Metric Embeddings, Impossibility and Information Geometry in Cartography

Distortion, Limits, and the Statistical Geometry of Map Projections

Silvano Salvador

Contents

Abstract vi

I Metric Embedding Theory and the Geometry of Distortion 1

1 Introduction 2

1.1 From an empirical index to three theoretical questions . . . 2

1.2 What is proved, and what is left open, in each part . . . 3

1.3 Relationship to the companion volume . . . 4

2 Preliminaries: Distortion, Doubling, and the Classical Embedding Theorems 5

2.1 Distortion and bi-Lipschitz embeddings . . . 5

2.2 Doubling metric spaces . . . 6

2.3 Bourgain's embedding theorem . . . 6

2.4 Assouad's embedding theorem . . . 6

2.5 The dimension gap: why neither theorem answers the cartographic question . 7

3 The Chordal Benchmark: An Explicit Distortion Bound for Small Spherical Caps 8

3.1 The chordal distance on the sphere . . . 8

3.2 An explicit planar embedding: orthographic projection . . . 9

3.3 The chordal benchmark theorem . . . 9

3.4 Numerical verification . . . 10

3.5 What remains open . . . 11

4 A Universal Lower Bound: Ptolemy's Inequality and the Limits of Any Embedding 12

4.1 Ptolemy's inequality in the Euclidean plane . . . 12

4.2 A symmetric spherical quadruple . . . 13

4.3 The spherical square violates Ptolemy's inequality . . . 13

4.4 The universal distortion lower bound . . . 14

4.5 What this chapter does, and does not, establish . . . 15

5 A Better Embedding: The Azimuthal Equidistant Projection 17

5.1 Motivation: orthographic projection is not the best embedding . . . 17

5.2 The azimuthal equidistant projection, restricted to a cap . . . 17
5.3 Radial pairs: exact isometry . . . 18
5.4 Non-radial pairs: the tangential stretch . . . 18
5.5 The global distortion of the azimuthal equidistant projection . . . 20
5.5.1 Asymptotic comparison . . . 20
5.6 What this chapter does, and does not, establish . . . 21

6 The Complete Azimuthal Family: Equal-Area and Stereographic Projections 23
6.1 A unifying reformulation of azimuthal projections . . . 23
6.1.1 Tangential pairs: a direct corollary of Chapter 5 . . . 23
6.1.2 Radial pairs: concavity and convexity . . . 24
6.2 The Lambert Azimuthal Equal-Area projection . . . 25
6.3 The stereographic projection . . . 26
6.4 The coincidence, and the four-projection comparison . . . 27
6.4.1 Asymptotic ordering . . . 27
6.5 What this chapter does, and does not, establish . . . 28

7 A General Curvature Theorem, the Gnomonic Projection, and the Local Optimality of the Equidistant Profile 30
7.1 A general theorem for profiles of fixed curvature sign . . . 30
7.2 Recovering the equal-area and stereographic results . . . 32
7.3 The gnomonic projection . . . 32
7.4 The five-projection comparison, completed . . . 33
7.4.1 Asymptotic ordering, extended . . . 33
7.5 Local optimality of the equidistant profile . . . 34
7.6 What this chapter does, and does not, establish . . . 35

8 A General First-Variation Theorem: The Equidistant Profile as a Gateaux-Stationary Point 37
8.1 Setup: an arbitrary smooth perturbation of the equidistant profile . . . 37
8.2 The supremum branch: a non-degenerate corner maximizer . . . 38
8.3 The infimum branch: a genuinely degenerate tie . . . 39
8.4 The general local-optimality theorem . . . 40
8.5 Numerical verification against the grid pipeline . . . 41
8.6 What this chapter does, and does not, establish . . . 42

9 Strict Gateaux Optimality: Closing the Degenerate-Direction Gap, and Kinked Perturbations 44

9.1 Degenerate directions do not exist among nontrivial perturbations 44
9.2 Extension to piecewise-C^1 (kinked) perturbations 45
9.3 Numerical verification 47
9.4 What this chapter closes, and what remains 48

10 Removing the Auxiliary Hypotheses: An Unconditional Curvature Theorem, and the Cubic Family in Closed Form at Every Radius 49
10.1 An unconditional curvature theorem 49
10.2 The cubic perturbation family, exactly, at every radius 50
10.3 Numerical verification 51
10.4 What this chapter closes, and what remains 52

11 Closing a Gap: The Orthographic Projection via the General Curvature Theorem 53
11.1 The orthographic profile is concave, with an exact boundary hypothesis 53
11.2 An exact theorem, superseding Chapter 3's numerically-supported proposition 54
11.3 Numerical corroboration, independent of Chapter 3's own check 54
11.4 What this chapter closes, and what it does not 55

12 Beyond Perturbations: Formalizing the Global Optimality Problem 56
12.1 The question, precisely restated 56
12.2 Three gaps, not one 57
12.3 Why the class gap resists rearrangement 57
12.4 What a resolution would require: a three-part programme 58
12.5 The existence question 59
12.6 What this chapter closes, and what remains 59

II An Impossibility Theorem for the Informational Distortion Index 61

13 The Informational Distortion Index, Revisited 62
13.1 Sampling regimes 62
13.2 Uniform minimizers and regime-dependent rankings 63
13.3 What Part II proves, and does not 63

14 An Impossibility Theorem for the Informational Distortion Index 64
14.1 A general reduction: one adversarial pair suffices 64
14.2 The main theorem 64
14.3 Why the ranking flips: a structural remark, not a proof 65
14.4 What remains open 66

15 Numerical Illustration: The Impossibility Theorem on Genuine Grids 67
15.1 Methodology and setup . . . 67
15.1.1 Sampling regimes . . . 67
15.1.2 Geodesic distances and projections . . . 67
15.2 Full results . . . 68
15.3 Reading the impossibility off the table . . . 68
15.4 Closing the roadmap for Part II . . . 69

16 The Impossibility Persists: A Larger Pool and a Fifth Regime 70
16.1 Setup: seventeen projections, five regimes . . . 70
16.2 Full results . . . 70
16.3 Reading the extended impossibility off the table . . . 70
16.4 What remains open . . . 71

III Fisher–Rao Geometry of Map Projections 73

17 The Fisher–Rao Geometry of the Categorical Model 74
17.1 The categorical model and its interior . . . 74
17.2 The Fisher information metric of the categorical model . . . 75
17.3 The square-root map and its isometry to a sphere . . . 75
17.4 The Fisher–Rao distance in closed form . . . 76
17.5 Constant curvature . . . 77
17.6 Local relation to the Jensen–Shannon divergence . . . 77
17.7 A global sandwich bound between the Fisher–Rao and Hellinger quantities . . . 78
17.8 What this chapter establishes, and what remains open . . . 78

18 Map Projections as Points on a Statistical Manifold 80
18.1 The embedding of a projection pool into the manifold . . . 80
18.2 The pairwise Fisher–Rao distance matrix, Global regime . . . 81
18.3 An exact coincidence: Mercator and Lambert Conformal Conic . . . 81
18.4 Explicit geodesics on the statistical manifold . . . 82
18.5 What a geodesic between two projections means, and does not mean . . . 83
18.6 What this chapter proves, and what remains open . . . 83

19 Numerical Illustration: Fisher–Rao Distances Across the Companion Volume's Pool 84
19.1 Methodology and setup . . . 84
19.2 Full results . . . 84

19.3 Cross-validating the impossibility theorem under a different metric 85
19.4 Where the two full rankings disagree 85
19.5 Confirming the global sandwich on genuine cartographic data 86
19.6 Closing the roadmap for Part III 87

20 Regional Pairwise Structure and Hierarchical Clustering of the Projection Pool 88
20.1 Method 88
20.2 Cophenetic fidelity by regime 88
20.3 Merge structure by regime 89
20.4 Linkage sensitivity 90
20.5 Cluster stability under grid resampling 90
20.5.1 Method 90
20.5.2 Results 91
20.6 What remains open 92

21 Synthesis: Three Lenses on One Phenomenon 93
21.1 What was proved, part by part 93
21.1.1 A stopping rule for Part I's azimuthal-profile sequence 93
21.2 Three unplanned echoes across the parts 95
21.3 What this monograph does not claim 96
21.4 Closing remark 96

References 97

Abstract

This monograph is an independent companion to *Original Mathematical Perspectives on Cartography*, not a sequel: it takes the Informational Distortion Index (IDI) developed there as a point of departure rather than an object of further empirical testing, and pursues three genuinely different theoretical directions opened, but not developed, by that empirical program. **Part I** recasts the classical impossibility of isometric map projections (Gauss's Theorema Egregium) in the quantitative language of metric embedding theory, using Bourgain's and Assouad's embedding theorems as background and developing an explicit, elementary lower bound on the distortion forced on any embedding of a small spherical cap into the Euclidean plane; a closing general theorem unifies the resulting family of azimuthal-projection distortion computations by the sign of the profile's curvature alone, adds the gnomonic projection as a further worked example, and proves that the azimuthal equidistant projection is a strict local minimum of distortion within the natural cubic-perturbation family of nearby profiles; a further chapter then generalizes this local-optimality result, via a Gateaux/Danskin first-variation argument, to every smooth perturbation direction, not only cubic ones; Part I closes not by attempting the remaining global-optimality question but by formalizing it, decomposing it into three independent gaps and stating explicitly what a genuine resolution would require. **Part II** states and proves a formal impossibility theorem for the IDI itself: no single projection can minimize the index uniformly across all sampling regimes and region shapes, a rigorous no-free-lunch result for distributional distortion measures. **Part III** develops the Fisher–Rao geometry of the statistical manifold underlying the IDI's construction, treating each projection's induced distance distribution as a point on a statistical manifold and studying the geometry — geodesics, curvature — of the space of projections under this metric. As in the companion volume, every numerical claim is the direct, executed output of a real computational pipeline; every theorem is proved in full, and the boundary between what is proved here and what remains open is stated explicitly rather than blurred.

Part I

Metric Embedding Theory and the Geometry of Distortion

Chapter 1

Introduction

1.1 From an empirical index to three theoretical questions

Original Mathematical Perspectives on Cartography proposed the Informational Distortion Index (IDI): a Jensen–Shannon divergence between the distribution of true geodesic distances on a sampled region of the sphere and the distribution of the corresponding planar distances under a given map projection. That index was constructed, computed, and stress-tested entirely empirically, across a growing pool of projections, sampling grids, and — in the toponymic and shape-descriptor chapters that followed — real gazetteer data. Three theoretical questions were raised along the way but deliberately left unaddressed, precisely because answering them was recognized at the time as requiring different mathematical machinery from the executed-pipeline empiricism of that volume:

1. **How large, quantitatively, must distortion be?** Gauss's Theorema Egregium (companion volume, §2) proves that no local isometry between a curved region of the sphere and the plane exists at all — an absolute, qualitative impossibility. It says nothing about *how much* distortion an optimal embedding of a given region must incur, nor how that minimal distortion scales with the size of the region. Metric embedding theory — the branch of mathematics concerned with exactly this kind of quantitative distortion question for general metric spaces — supplies both a vocabulary and, in its classical theorems (Bourgain, Assouad), general-purpose machinery that has never been brought to bear specifically on the sphere-to-plane cartographic question. **Part I** of this monograph does so.

2. **Can the IDI itself be beaten uniformly?** The companion volume's empirical rankings (twelve, then seventeen projections, four and then five sampling grids) consistently found that the best-performing projection depends on the sampling regime: Aitoff wins globally but is the worst performer at mid-latitude; Hammer overtakes it at some bin counts; polyconic and conic constructions dominate near the pole while failing badly elsewhere. This pattern — no projection uniformly best — was observed repeatedly but never proved as a general, regime-independent fact about the IDI's mathematical structure. **Part II** states and proves such a theorem.

3. **Is there a natural geometry on the space of projections itself?** The IDI compares two *distributions* (geodesic and projected distance histograms). Any such comparison of probability distributions has a natural home in information geometry, where the Fisher–Rao metric turns a family of distributions into a genuine Riemannian manifold, with its own geodesics and curvature, of which the Jensen–Shannon divergence used throughout the companion volume is only a first-order, non-Riemannian approximation. **Part III** develops this geometry for the specific statistical manifold generated by map projections.

These three questions are not sequential steps toward a single unified theorem; they are three genuinely independent theoretical lenses on the same underlying phenomenon, in exactly the sense that Chapter 15 of the companion volume, on the toponymic/shape/elevation program,

found four empirical measures of national heterogeneity to be substantially uncorrelated with one another. The metric-embedding question concerns worst-case pointwise distortion; the impossibility question concerns the combinatorics of sampling regimes; the Fisher–Rao question concerns the differential geometry of a space of distributions. A reader interested in only one of the three parts can read it independently of the other two.

1.2 What is proved, and what is left open, in each part

Consistent with the methodological discipline of the companion volume, this monograph states plainly, at the outset, the exact boundary of what each part establishes.

Part I proves an explicit, elementary, and fully rigorous asymptotic lower bound (Theorem 3.3) on the distortion of one specific, natural embedding of a small spherical cap into the plane — the chordal (ambient Euclidean) embedding — as a function of the cap's geodesic radius and the sphere's curvature. This is presented honestly as a *benchmark* result, not yet the sharp lower bound over *all* possible embeddings of the cap into the plane; Chapter 3 closes by stating precisely what additional comparison-geometry machinery (Rauch's comparison theorem, or a Toponogov-type argument) would be needed to upgrade the benchmark into a fully general sharp bound, and this general bound is left as the explicit subject of a subsequent chapter rather than asserted prematurely. Chapter 4 supplies that general (embedding-independent) lower bound via Ptolemy's inequality, and Chapter 5 narrows the resulting sandwich further from the other side, exhibiting an explicit alternative embedding (azimuthal equidistant projection) whose distortion is strictly smaller than the chordal benchmark's; as that chapter states openly, this is a substitution for the originally planned Assouad-type sharpness construction, not a completion of it, and a nonzero gap remains even after this narrowing.

Part II's impossibility theorem, once developed, will be a genuine no-free-lunch theorem in the precise sense of that term in optimization theory: a proof that a specific family of adversarial sampling-regime pairs forces any projection to perform poorly on at least one member of the pair, however well it performs on the other. Its numerical illustration will follow the companion volume's discipline of executing real code against real sampling grids rather than asserting the qualitative pattern by appeal to the already-published empirical rankings alone. Chapter 16 extends this numerical illustration to the companion volume's own later seventeen-projection pool and a fifth (Arctic) regime, directly addressing one of the two open items Chapter 14's own closing section identifies.

Part III's Fisher–Rao geometry, once developed, will require care in one specific respect flagged here in advance: the geodesic-distance and projected-distance histograms used throughout the companion volume are discrete, finite-bin approximations to continuous densities, and the Fisher–Rao metric's standard closed forms (available for many parametric families) do not automatically apply to arbitrary empirical histograms without an explicit choice of underlying statistical model or a discrete (categorical-distribution) version of the Fisher–Rao metric. This choice will be made explicit, not silently assumed, when Part III is developed. Chapter 20 extends Part III's own pairwise Fisher–Rao analysis from the Global regime alone to all four regimes of this monograph's pipeline, and adds a hierarchical clustering analysis explicitly flagged as open at the end of Chapter 18.

1.3 Relationship to the companion volume

This monograph reuses three pieces of notation and one dataset established in the companion volume, and no others: the Informational Distortion Index itself (Chapter 8 of the companion volume), the haversine geodesic-distance and Lambert Azimuthal Equal-Area projection formulas used throughout its computational pipeline, and, where Part II's numerical illustration requires it, the same twelve-to-seventeen-projection pool and multi-grid sampling machinery. No result of the companion volume is assumed to be extended, corrected, or superseded by anything proved here; the two monographs are independent and can be read in either order, though the motivation in Section 1.1 above will read more naturally to a reader already familiar with the empirical IDI construction.

Chapter 2

Preliminaries: Distortion, Doubling, and the Classical Embedding Theorems

This chapter fixes notation and states, without proof, the two classical theorems of metric embedding theory — due to Bourgain and to Assouad — that frame the cartographic question of Chapter 3. Both theorems are general-purpose results about embedding arbitrary metric spaces into Euclidean space; neither was formulated with cartography in mind, and Section 2.5 explains precisely why neither can be applied off the shelf to answer the sphere-to-*plane* question this monograph is actually interested in, which is exactly the gap Chapter 3 begins to close.

2.1 Distortion and bi-Lipschitz embeddings

Definition 2.1 (Bi-Lipschitz embedding and distortion)**.** Let (X, d_X) and (Y, d_Y) be metric spaces and $f : X \to Y$ an injective map. The map f is a *bi-Lipschitz embedding with distortion* $D \geq 1$ if there exists a constant $c > 0$ such that

$$c\, d_X(x, y) \;\leq\; d_Y(f(x), f(y)) \;\leq\; cD\, d_X(x, y) \qquad \text{for all } x, y \in X. \tag{2.1}$$

Equivalently, writing $\text{expansion}(f) = \sup_{x \neq y} d_Y(f(x), f(y))/d_X(x, y)$ and $\text{contraction}(f) = \sup_{x \neq y} d_X(x, y)/d_Y(f(x), f(y))$, the distortion of f is $\text{dist}(f) = \text{expansion}(f) \cdot \text{contraction}(f)$, and D in the displayed inequality can always be taken equal to $\text{dist}(f)$ after rescaling f by the constant $c = 1/\text{contraction}(f)$. $D = 1$ recovers an exact scaled isometry.

This is the standard notion of distortion used throughout the metric embedding literature (Linial, London & Rabinovich, 1995) and is the one this monograph adopts. It is a purely metric notion: it makes no reference to angles, area, or any differential structure, unlike the classical pointwise cartographic distortion measures of the companion volume (Tissot's indicatrix, scale factors h, k). Remark 2.2 below makes the relationship between the two notions precise.

Remark 2.2 (Relation to Tissot's indicatrix). For a smooth projection f, restricted to a small enough neighbourhood of a point p, the local expansion and contraction factors of Definition 2.1 converge, as the neighbourhood shrinks to p, to the maximum and minimum singular values of Df_p — exactly Tissot's indicatrix semi-axes $h(p), k(p)$ from companion-volume §2.3. Distortion in the sense of Definition 2.1, applied to a small neighbourhood of p, is thus the pointwise ratio $h(p)/k(p)$ in the small-neighbourhood limit; the two frameworks are compatible, differing only in that Definition 2.1's distortion is inherently a global, two-point quantity (it can be evaluated exactly between any two points x, y, however far apart, without any differentiability assumption on f), while Tissot's indicatrix is inherently local and requires f to be differentiable at p.

2.2 Doubling metric spaces

Definition 2.3 (Doubling constant and doubling dimension)**.** A metric space (X, d) has *doubling constant* λ if every closed ball $B(x, r) \subset X$ can be covered by at most λ closed balls of radius $r/2$. The *doubling dimension* of X is $\log_2 \lambda$.

Euclidean space $\mathbb{R}^n$ has doubling dimension exactly n (up to a dimension-dependent constant in λ). Any compact Riemannian manifold is doubling, with a doubling constant controlled by its dimension and by bounds on its sectional curvature and injectivity radius (a standard consequence of the Bishop–Gromov volume comparison theorem); in particular, the round sphere S^2_R of any radius R, and every geodesically convex subset of it, is a doubling metric space with doubling dimension 2, matching its topological and Hausdorff dimension. This single fact is what makes Assouad's theorem (Theorem 2.5 below) relevant to the cartographic setting at all.

2.3 Bourgain's embedding theorem

Theorem 2.4 (Bourgain, 1985)**.** *There is an absolute constant C such that every finite metric space (X, d) with $|X| = n$ points admits a bi-Lipschitz embedding into Euclidean space (of some dimension depending on n) with distortion at most $C \log n$.*

Bourgain's theorem is best possible in general: Linial, London and Rabinovich (1995) showed that the shortest-path metric of a constant-degree expander graph on n vertices requires distortion $\Omega(\log n)$ in any embedding into Euclidean space, matching Theorem 2.4's upper bound up to the constant C. Theorem 2.4 makes no assumption whatsoever on the structure of X — it applies equally to the discrete metric on n arbitrary points sampled from the sphere, from an expander graph, or from any other finite metric space — and correspondingly gives no better bound for the sphere than for a worst-case metric space of the same cardinality. It is stated here purely as background: nothing in this monograph relies on Theorem 2.4 directly, precisely because its embedding target has unrestricted dimension (Section 2.5), but it is the natural starting point for any reader's intuition about how distortion scales with the number of points involved, and it explains why the finite-sample, n-dependent embedding results that would follow from applying it naively to the companion volume's population-weighted country samples are not, in fact, useful for the plane-specific question this monograph asks.

2.4 Assouad's embedding theorem

Theorem 2.5 (Assouad, 1983)**.** *Let (X, d) be a metric space with doubling dimension at most k, and fix $0 < \varepsilon < 1$. The* snowflaked *metric space (X, d^ε) — the same set X equipped with the metric $d^\varepsilon(x, y) := d(x, y)^\varepsilon$ — admits a bi-Lipschitz embedding into some Euclidean space $\mathbb{R}^N$, with both the target dimension N and the distortion bounded by constants depending only on k and ε (not on $|X|$, and not requiring X to be finite).*

Two features of Theorem 2.5 make it, at first glance, far better suited to the cartographic question than Theorem 2.4: the distortion and target dimension bounds are dimension-dependent constants, uniform over the (possibly infinite, possibly continuous) space X, with no $\log n$ blow-up as more points are added, and the doubling-dimension-2 sphere is exactly the kind of space Theorem 2.5 was built for. Section 2.5 explains why this promise is nonetheless not directly usable here.

2.5 The dimension gap: why neither theorem answers the cartographic question

Why Chapter 3 is necessary

Both Theorems 2.4 and 2.5 embed a metric space into *some* Euclidean space, of dimension left unspecified (Bourgain) or bounded only by a constant depending on the doubling dimension and ε, which for the sphere's doubling dimension 2 is not generally 2 itself (Assouad). Neither theorem, as stated, guarantees an embedding into $\mathbb{R}^2$ specifically — and a map projection is, by definition, exactly a map into $\mathbb{R}^2$. Moreover, Theorem 2.5 does not embed (X, d) itself but the *snowflaked* metric (X, d^ε) for some $\varepsilon < 1$: since $d^\varepsilon \neq d$ for $\varepsilon < 1$, an Assouad embedding of the snowflake is not a low-distortion embedding of the original sphere metric, and offers no direct information about the distortion of any map projection, which must respect the original geodesic metric d, not a snowflaked version of it. Both classical theorems, in short, are embedding existence results for essentially unconstrained target dimension and, in Assouad's case, a modified metric; cartography imposes exactly the constraint — target dimension identically 2, original (non-snowflaked) metric — that both theorems leave unaddressed. This is precisely the gap Chapter 3 begins to close, not by invoking either classical theorem directly, but by an elementary, self-contained argument specific to the sphere and the plane.

Chapter 3

The Chordal Benchmark: An Explicit Distortion Bound for Small Spherical Caps

Chapter 2 explained why neither Bourgain's nor Assouad's theorem, as stated, answers the specific question this monograph asks: how much distortion is forced on an embedding of a small geodesic cap of the sphere into the *plane*, specifically. This chapter answers that question for one natural, explicit embedding — the orthographic projection — rather than for the class of all possible embeddings into $\mathbb{R}^2$. As flagged already in Section 1.2, the result of this chapter is a *benchmark*: a fully rigorous distortion bound for one specific map, not yet the sharp lower bound over every embedding, which is deferred to a subsequent chapter.

3.1 The chordal distance on the sphere

Fix a sphere S^2_R of radius R and two points $p, q \in S^2_R$ separated by geodesic (great-circle) distance $\ell = d_{S^2_R}(p, q)$. Embedded in the ambient space $\mathbb{R}^3$, p and q also have a straight-line Euclidean distance, the *chordal distance*:

$$c(\ell) \;=\; 2R \sin\!\left(\frac{\ell}{2R}\right). \tag{3.1}$$

This is elementary: if $\varphi = \ell/R$ is the angle subtended at the sphere's centre by p and q, the chord pq is the base of an isosceles triangle with two sides of length R and apex angle φ, giving $c = 2R\sin(\varphi/2)$.

Remark 3.1 (Chordal distance is not a map projection). Equation (3.1) embeds S^2_R isometrically into its own ambient $\mathbb{R}^3$ up to the (exact, closed-form) distortion of replacing arc length by chord length; it is not, by itself, an embedding into $\mathbb{R}^2$, and so does not answer the cartographic question directly. Its role in this chapter is different: since $c(\ell) \le \ell$ always (a chord is never longer than the arc it subtends), and since orthogonal projection from $\mathbb{R}^3$ onto any plane is 1-Lipschitz (it never increases Euclidean distance), the chordal distance gives an upper bound on the planar distance produced by *any* projection of S^2_R obtained by first embedding into $\mathbb{R}^3$ and then projecting orthogonally onto a plane — which is exactly how Section 3.2 constructs its explicit $\mathbb{R}^2$-valued embedding.

Expanding Eq. (3.1) for small ℓ/R,

$$c(\ell) \;=\; \ell - \frac{\ell^3}{24R^2} + O\!\left(\frac{\ell^5}{R^4}\right), \tag{3.2}$$

so the chord contracts relative to the arc by a relative amount $\ell^2/(24R^2)$ to leading order. This is the correct order of magnitude for *any* distortion measure tied to the sphere's curvature $1/R^2$: it is the same quadratic-in-curvature scaling that appears, from a completely different argument,

in the classical proof of Theorema Egregium (companion volume, §2), and it reappears, with a different constant, in the orthographic-projection bound of Theorem 3.3 below.

3.2 An explicit planar embedding: orthographic projection

To turn Remark 3.1's observation into a genuine cartographic statement, fix a point $p_0 \in S^2_R$ and use geodesic polar coordinates (φ, θ) centred at p_0: $\varphi \in [0, \pi]$ is the colatitude (geodesic distance from p_0 is $R\varphi$) and $\theta \in [0, 2\pi)$ is the azimuth. The *orthographic projection* centred at p_0 sends the point at (φ, θ) to the planar point

$$\Pi(\varphi, \theta) \;=\; \big(R \sin\varphi \cos\theta,\; R \sin\varphi \sin\theta\big) \;\in\; \mathbb{R}^2, \tag{3.3}$$

i.e., the point obtained by embedding S^2_R in $\mathbb{R}^3$ in the usual way with p_0 at the "north pole" and orthogonally projecting onto the tangent plane at p_0. This is precisely the composition described in Remark 3.1: embed, then project orthogonally.

Lemma 3.2 (Exact local scale factors of Π)**.** *At every point of colatitude $\varphi \in (0, \pi/2)$, the differential of Π has singular values 1 (in the transverse, i.e., constant-φ direction) and $\cos\varphi$ (in the radial, i.e., constant-θ direction).*

Proof. The metric on S^2_R in geodesic polar coordinates is $R^2(\mathrm{d}\varphi^2 + \sin^2\varphi\, \mathrm{d}\theta^2)$, so an infinitesimal transverse displacement $R \sin\varphi\, \mathrm{d}\theta$ and an infinitesimal radial displacement $R\, \mathrm{d}\varphi$ are the two orthogonal principal directions at (φ, θ). Differentiating Eq. (3.3): in the transverse direction, the image displacement has length $R \sin\varphi\, \mathrm{d}\theta$ exactly, giving scale factor 1; in the radial direction, $\mathrm{d}(R \sin\varphi) = R \cos\varphi\, \mathrm{d}\varphi$, giving scale factor $\cos\varphi$. □

This is the Tissot indicatrix of the orthographic projection in the language of Remark 2.2: circles of constant colatitude are mapped without distortion ($h = 1$), while radial segments are contracted by $\cos\varphi$ ($k = \cos\varphi < 1$ for $\varphi > 0$), so the indicatrix is an ellipse with semi-axes 1 and $\cos\varphi$, degenerating to a segment as $\varphi \to \pi/2$ (the projection folds over at the equator, which is why only caps with $\varphi_{\max} < \pi/2$ are considered below).

3.3 The chordal benchmark theorem

Let $B(p_0, r) \subset S^2_R$ denote the closed geodesic cap of radius $r < \pi R/2$ centred at p_0, i.e., the set of points at colatitude $\varphi \le r/R$.

Theorem 3.3 (Chordal benchmark)**.** *Restricted to $B(p_0, r)$, the orthographic projection Π of Eq. (3.3) satisfies, for all $x, y \in B(p_0, r)$,*

$$|\Pi(x) - \Pi(y)| \;\le\; d_{S^2_R}(x, y), \tag{3.4}$$

with equality approached in the limit $x \to y$ along a transverse direction; and, restricted to any radial *pair of points (equal azimuth θ, colatitudes $\varphi_1 \neq \varphi_2 \le r/R$),*

$$d_{S^2_R}(x, y) \;\le\; \sec\!\Big(\frac{r}{R}\Big)\, |\Pi(x) - \Pi(y)|. \tag{3.5}$$

Consequently, the restriction of Π to radial segments through p_0 is a bi-Lipschitz embedding of distortion exactly $\sec(r/R)$, and for small r/R,

$$\sec\!\Big(\frac{r}{R}\Big) \;=\; 1 + \frac{1}{2}\Big(\frac{r}{R}\Big)^2 + O\!\left(\Big(\frac{r}{R}\Big)^4\right), \tag{3.6}$$

matching the quadratic-in-curvature scaling of Eq. (3.2) and of Theorema Egregium.

Proof. Expansion bound (3.4). By Remark 3.1, $|\Pi(x) - \Pi(y)| \le c(d_{S^2_R}(x,y)) \le d_{S^2_R}(x,y)$, the first inequality because orthogonal projection onto the tangent plane at p_0 is 1-Lipschitz on the ambient chordal distance, the second because a chord is never longer than the arc it subtends. Equality in the limit follows from Lemma 3.2: along the transverse direction the scale factor is exactly 1 at every colatitude, so the ratio $|\Pi(x) - \Pi(y)|/d_{S^2_R}(x,y) \to 1$ as x, y approach each other transversally.

Radial contraction bound (3.5). For a radial pair at fixed azimuth θ, Eq. (3.3) gives $|\Pi(x) - \Pi(y)| = R|\sin\varphi_2 - \sin\varphi_1|$, while $d_{S^2_R}(x,y) = R|\varphi_2 - \varphi_1|$. By the mean value theorem applied to sin, there is some ξ strictly between φ_1 and φ_2 with $\sin\varphi_2 - \sin\varphi_1 = (\varphi_2 - \varphi_1)\cos\xi$, so

$$\frac{d_{S^2_R}(x,y)}{|\Pi(x) - \Pi(y)|} = \frac{1}{\cos\xi} \le \frac{1}{\cos(r/R)} = \sec\left(\frac{r}{R}\right),$$

since $\xi \le \max(\varphi_1, \varphi_2) \le r/R$ and cos is decreasing on $[0, \pi/2)$. This bound is attained in the limit $\varphi_1, \varphi_2 \to r/R$, so $\sec(r/R)$ is the exact (not merely an upper) distortion constant for radial pairs.

The asymptotic expansion (3.6) is the standard Taylor series of sec at 0. □

What this theorem does, and does not, establish

Theorem 3.3 is stated and proved only for *radial* pairs of points, i.e., pairs lying on a common meridian through the cap centre p_0; this is exactly the direction in which Lemma 3.2 identifies the projection's maximal contraction, and the mean-value-theorem argument above is fully rigorous and elementary. It does *not*, as it stands, prove that $\sec(r/R)$ bounds the distortion of *every* pair $x, y \in B(p_0, r)$, radial or not: a pair separated by both a colatitude difference and an azimuthal difference is not covered by the proof above, and asserting the same bound for such pairs without proof would violate this monograph's discipline of not stating as proved what has only been checked numerically. Section 3.4 reports an extensive numerical check, across many random non-radial pairs and cap radii, that the bound (3.5) in fact holds for *all* pairs in $B(p_0, r)$, not only radial ones — consistent with radial pairs being the extremal (worst-case) configuration, as the Tissot-ellipse structure of Lemma 3.2 suggests. But this general two-point statement is reported here as a numerically well-supported *proposition*, not a theorem: a fully symbolic proof would require controlling how the ratio $d_{S^2_R}(x,y)/|\Pi(x) - \Pi(y)|$ varies with the azimuthal separation $\Delta\theta$ between x and y, which is a genuinely different (and harder) argument from the one-dimensional mean-value-theorem argument used above, and is not attempted here. This is a strictly narrower gap than the one left open for Chapter 4: that chapter's open problem is the sharp lower bound over *all possible embeddings* into $\mathbb{R}^2$, not merely the general two-point case of this one specific embedding.

3.4 Numerical verification

Before reporting the general two-point bound as a proposition, it was checked numerically, on the unit sphere ($R = 1$), as follows. For each of seven cap radii $r \in \{0.02, 0.05, 0.1, 0.2, 0.3, 0.5, 0.7\}$ (radians), 4,000 points were sampled with uniform area density inside the cap $B(p_0, r)$, and 200,000 random ordered pairs of distinct points were drawn from this sample (pairs with negligible geodesic separation,

$d_{S^2_R}(x,y) < 10^{-9}$, were discarded to avoid division by near-zero). For each pair the ratio $|\Pi(x) - \Pi(y)|/d_{S^2_R}(x,y)$ was computed exactly from Eq. (3.3) and the exact great-circle formula $d_{S^2_R}(x,y) = R\arccos\big(\cos\varphi_x\cos\varphi_y + \sin\varphi_x\sin\varphi_y\cos(\theta_x - \theta_y)\big)$; the maximum and minimum of this ratio over all sampled pairs give an empirical bi-Lipschitz distortion $\mathrm{dist}_{\mathrm{emp}} = \text{max ratio}/\text{min ratio}$, directly comparable to the theoretical bound $\sec(r)$ of Theorem 3.3.

Table 3.1: Empirical distortion of the orthographic projection over 200,000 random point pairs per cap radius, unit sphere ($R = 1$), against the theoretical radial bound $\sec(r)$.

r (rad)	max ratio	min ratio	$\mathrm{dist}_{\mathrm{emp}}$	$\sec(r)$	$\mathrm{dist}_{\mathrm{emp}} \le \sec(r)$
0.02	1.000000	0.999807	1.000193	1.000200	yes
0.05	1.000000	0.998796	1.001205	1.001251	yes
0.10	0.999999	0.995403	1.004618	1.005021	yes
0.20	1.000000	0.981022	1.019345	1.020339	yes
0.30	0.999997	0.956961	1.044972	1.046752	yes
0.50	0.999996	0.882869	1.132667	1.139494	yes
0.70	0.999991	0.779895	1.282213	1.307459	yes

Three features of Table 3.1 confirm the theoretical picture of Theorem 3.3 and Lemma 3.2 rather than merely being consistent with it. First, the maximum ratio is 1 to within sampling resolution at every cap radius, exactly as Eq. (3.4) predicts (transverse pairs are never expanded, and near-coincident transverse pairs approach ratio 1). Second, the minimum ratio decreases monotonically with r and tracks $\cos(r)$ closely (e.g., at $r = 0.5$, $\cos(0.5) = 0.877583$ against an empirical minimum of 0.882869, the small excess being the expected finite-sample effect of not sampling exactly at the cap boundary). Third, and most directly relevant to the disclaimer above, the empirical distortion $\mathrm{dist}_{\mathrm{emp}}$ — computed over *all* sampled pairs, radial and non-radial alike — never exceeds the theoretical radial bound $\sec(r)$ at any of the seven radii tested, supporting the proposition that radial pairs are indeed the extremal configuration for this projection, without yet constituting a proof of it.

3.5 What remains open

This chapter has proved, with an elementary and fully rigorous argument, that the orthographic projection of a geodesic cap of radius r is bi-Lipschitz on radial segments with distortion exactly $\sec(r/R)$, matching to leading order the quadratic-in-curvature scale already visible in the chordal-distance expansion of Eq. (3.2). Two extensions are left open, of genuinely different character, and are recorded here rather than blurred into the theorem above:

1. **Extending Theorem 3.3 to all pairs, not just radial ones.** Section 3.4's numerics support this extension strongly but do not prove it; a proof would control the ratio $d_{S^2_R}(x,y)/|\Pi(x) - \Pi(y)|$ as a function of azimuthal separation $\Delta\theta$, showing it is maximised at $\Delta\theta = 0$ (the radial case already proved).

2. **The sharp lower bound over all embeddings, not just Π.** Theorem 3.3 bounds the distortion of one specific, natural embedding. It says nothing about whether some other embedding of $B(p_0, r)$ into $\mathbb{R}^2$ — not of the orthographic form (3.3) — might achieve strictly smaller distortion. Settling this requires comparison-geometry machinery (a Rauch- or Toponogov-type argument) genuinely absent from this chapter's elementary toolkit, and is the explicit subject of the roadmap chapter that follows.

Chapter 4

A Universal Lower Bound: Ptolemy's Inequality and the Limits of Any Embedding

Chapter 3 proved a distortion bound for one specific, natural embedding of a geodesic cap into the plane — orthographic projection. It left open, honestly, whether some other embedding might do better. This chapter closes part of that gap: it proves a lower bound on the distortion of *every* bi-Lipschitz embedding of a suitable spherical configuration into $\mathbb{R}^2$, using a classical comparison-geometry fact — Ptolemy's inequality — rather than any property specific to orthographic projection. As Section 4.5 makes explicit, this lower bound is not shown to be sharp: a gap remains between it and Chapter 3's upper bound, and closing that gap is left as an open problem rather than asserted away.

Why Ptolemy's inequality, not the general Toponogov theorem

The roadmap for this chapter originally anticipated invoking Toponogov's comparison theorem, which bounds the angles of a geodesic triangle in a manifold of curvature *at least* κ below by the angles of the corresponding triangle in the constant-curvature-κ model space. That generality is not needed here: the round sphere S^2_R does not merely have curvature *bounded below* by $1/R^2$, it has curvature *exactly* $1/R^2$ everywhere, so every comparison this chapter needs reduces to elementary, exact spherical trigonometry (the spherical law of cosines) rather than the inequality-form machinery Toponogov's theorem provides for non-constant-curvature manifolds. The genuinely load-bearing comparison-geometry fact used below is a different, equally classical one: Ptolemy's inequality holds in the Euclidean plane (indeed in every space of non-positive curvature; see Foertsch, Lytchak & Schroeder, *Non-positive curvature and the Ptolemy inequality*, IMRN 2007) and can fail in spaces of positive curvature such as the sphere. This is the actual mechanism of the impossibility proved below, and it is elementary enough to prove from scratch in Section 4.1, without appeal to the general theory of either theorem.

4.1 Ptolemy's inequality in the Euclidean plane

Proposition 4.1 (Ptolemy's inequality)**.** *For any four points $A, B, C, D \in \mathbb{R}^2$,*

$$|AC| \cdot |BD| \;\le\; |AB| \cdot |CD| \;+\; |AD| \cdot |BC|. \tag{4.1}$$

Proof. Identify $\mathbb{R}^2$ with $\mathbb{C}$ and let $z_1, z_2, z_3, z_4 \in \mathbb{C}$ be the images of A, B, C, D. The algebraic identity

$$(z_1 - z_2)(z_3 - z_4) \;+\; (z_1 - z_4)(z_2 - z_3) \;=\; (z_1 - z_3)(z_2 - z_4) \tag{4.2}$$

holds for arbitrary complex numbers (expand both sides; every term cancels or matches). Taking the modulus of Eq. (4.2) and applying the triangle inequality $|u + v| \le |u| + |v|$ to its

left-hand side gives

$$|z_1-z_3|\,|z_2-z_4| \;=\; \big|(z_1-z_2)(z_3-z_4)+(z_1-z_4)(z_2-z_3)\big| \;\le\; |z_1-z_2|\,|z_3-z_4|+|z_1-z_4|\,|z_2-z_3|,$$

which is exactly Eq. (4.1) with A, B, C, D relabelled as z_1, z_2, z_3, z_4. □

Proposition 4.1 is purely a fact about $\mathbb{R}^2$ (equivalently $\mathbb{C}$): it makes no reference to the sphere and holds for any four points whatsoever, in any configuration. The strategy of this chapter is to exhibit four points on the sphere for which the *spherical* analogue of Eq. (4.1) fails, and to show that this failure forces a quantitative distortion penalty on any map sending those four points into the plane, however that map is constructed.

4.2 A symmetric spherical quadruple

Fix a cap centre $p_0 \in S^2_R$ and a colatitude $x = r/R \in (0, \pi/2)$, and let A, B, C, D be the four points at colatitude x from p_0 and azimuths $0°, 90°, 180°, 270°$ respectively — the vertices, in cyclic order, of a square inscribed in the boundary circle of the geodesic cap $B(p_0, r)$. By the spherical law of cosines, two points at common colatitude x separated by azimuth $\Delta\theta$ have geodesic distance $R \arccos\big(\cos^2 x + \sin^2 x \cos\Delta\theta\big)$. Applying this to adjacent vertices ($\Delta\theta = 90°$, $\cos\Delta\theta = 0$) and to diagonal vertices ($\Delta\theta = 180°$, $\cos\Delta\theta = -1$):

$$s := d_{S^2_R}(A,B) = d_{S^2_R}(B,C) = d_{S^2_R}(C,D) = d_{S^2_R}(D,A) = R\arccos(\cos^2 x), \tag{4.3}$$

$$d := d_{S^2_R}(A,C) = d_{S^2_R}(B,D) = R\arccos(\cos 2x) = 2Rx, \tag{4.4}$$

the second equality in Eq. (4.4) because A, p_0, C are colinear along a geodesic (diametrically opposite points on the boundary circle, joined through the centre), so $d(A, C)$ is simply twice the cap radius, $2Rx$; this is also consistent with $\arccos(\cos 2x) = 2x$ for $2x \in [0, \pi]$, which holds throughout the admissible range $x < \pi/2$.

By construction, A, B, C, D all lie in the closed cap $B(p_0, r)$ (in fact on its boundary), so any embedding $f : B(p_0, r) \to \mathbb{R}^2$ restricts to a map of these four points, to which Proposition 4.1 applies to the images.

4.3 The spherical square violates Ptolemy's inequality

If the sphere behaved like the Euclidean plane, an isometric copy of A, B, C, D would satisfy Eq. (4.1) with equality (a square is a cyclic quadrilateral, the equality case of Ptolemy's theorem): $d \cdot d = s \cdot s + s \cdot s$, i.e., $d = s\sqrt{2}$. Theorem 4.2 shows this fails on the sphere, strictly, in the direction of positive curvature.

Theorem 4.2 (Failure of the Euclidean diagonal relation)**.** *For every $x \in (0, \pi/2)$, the side length s and diagonal length d of Eqs. (4.3) and (4.4) satisfy*

$$d \;>\; s\sqrt{2}. \tag{4.5}$$

Proof. Define $g(x) = 2x - \sqrt{2}\arccos(\cos^2 x)$ for $x \in [0, \pi/2)$, so that Eq. (4.5) (recalling $d = 2Rx$ and $s = R\arccos(\cos^2 x)$) is equivalent to $g(x) > 0$. Clearly $g(0) = 0 - \sqrt{2}\arccos(1) = 0$. Differentiating, with $u(x) = \cos^2 x$,

$$\frac{\mathrm{d}}{\mathrm{d}x}\arccos(u(x)) \;=\; \frac{-u'(x)}{\sqrt{1-u(x)^2}} \;=\; \frac{\sin(2x)}{\sqrt{1-\cos^4 x}} \;=\; \frac{2\sin x\cos x}{\sin x\sqrt{1+\cos^2 x}} \;=\; \frac{2\cos x}{\sqrt{1+\cos^2 x}},$$

using $1-\cos^4 x = (1-\cos^2 x)(1+\cos^2 x) = \sin^2 x(1+\cos^2 x)$ and $\sin x > 0$ on $(0,\pi/2)$. Hence

$$g'(x) \;=\; 2 - \sqrt{2}\cdot\frac{2\cos x}{\sqrt{1+\cos^2 x}} \;=\; 2\left(1 - \frac{\sqrt{2}\cos x}{\sqrt{1+\cos^2 x}}\right).$$

Since $\cos x \in (0,1)$ on $(0,\pi/2)$, we have $\cos^2 x < 1$, so $1+\cos^2 x > 2\cos^2 x$, so $\sqrt{1+\cos^2 x} > \sqrt{2}\cos x$, so the fraction above is strictly less than 1, so $g'(x) > 0$ throughout $(0,\pi/2)$. Since $g(0) = 0$ and g is strictly increasing on $(0,\pi/2)$, $g(x) > 0$ for all x in that range, which is Eq. (4.5). □

By Theorem 4.2, $d^2 > 2s^2$, i.e., $d\cdot d > s\cdot s + s\cdot s$: the spherical square's diagonals and sides violate Eq. (4.1) in exactly the direction Ptolemy's inequality forbids for points in $\mathbb{R}^2$. Table 4.1 below reports exact evaluations of s and d at several cap radii, confirming the violation concretely rather than leaving Theorem 4.2 as a purely symbolic statement.

4.4 The universal distortion lower bound

Theorem 4.3 (Universal lower bound via Ptolemy's inequality)**.** *Let $f : B(p_0,r) \to \mathbb{R}^2$ be any injective map with distortion D in the sense of Definition 2.1 (i.e., there is $c > 0$ with $c\, d_{S^2_R}(x,y) \le |f(x)-f(y)| \le cD\, d_{S^2_R}(x,y)$ for all $x, y \in B(p_0,r)$). Then*

$$D \;\ge\; \frac{d}{s\sqrt{2}} \;=\; \frac{\sqrt{2}\,x}{\arccos(\cos^2 x)}, \qquad x = \frac{r}{R}. \tag{4.6}$$

Proof. Apply Proposition 4.1 to the images $f(A), f(B), f(C), f(D)$ of the spherical square of Section 4.2:

$$|f(A)-f(C)|\cdot|f(B)-f(D)| \;\le\; |f(A)-f(B)|\cdot|f(C)-f(D)| + |f(A)-f(D)|\cdot|f(B)-f(C)|.$$

By the definition of D, each diagonal image distance satisfies $|f(A)-f(C)|, |f(B)-f(D)| \ge c\,d$, and each side image distance satisfies $|f(A)-f(B)|, |f(B)-f(C)|, |f(C)-f(D)|, |f(D)-f(A)| \le cD\,s$. Substituting these bounds (lower bound on the left-hand side, upper bound on the right-hand side of the displayed inequality, both legitimate directions since all quantities are non-negative):

$$c^2 d^2 \;\le\; (cDs)(cDs) + (cDs)(cDs) \;=\; 2c^2D^2s^2.$$

Dividing by $c^2 s^2 > 0$ gives $d^2/s^2 \le 2D^2$, i.e., $D \ge d/(s\sqrt{2})$, which is Eq. (4.6) after substituting Eqs. (4.3) and (4.4). □

Key Result

Theorem 4.3 bounds the distortion of *every* embedding of $B(p_0,r)$ into $\mathbb{R}^2$ — not merely the orthographic projection of Chapter 3 — using nothing beyond the exact spherical law of cosines and the elementary Euclidean fact of Proposition 4.1. Theorem 4.2 guarantees the right-hand side of Eq. (4.6) exceeds 1 for every $x \in (0,\pi/2)$, so no embedding of any spherical cap, however cleverly constructed, can be a perfect isometry ($D = 1$); this recovers, quantitatively rather than merely qualitatively, the impossibility content of Theorema Egregium.

For small $x = r/R$, expanding Eq. (4.6) (using $\arccos(\cos^2 x) = \sqrt{2}\,x\,(1 - x^2/12 + O(x^4))$, obtained from the standard expansion $\arccos(1-u) = \sqrt{2u}\,(1 + u/12 + O(u^2))$ with $u = x^2 - x^4/3 + O(x^6)$):

$$D_{\min}(x) \;:=\; \frac{\sqrt{2}\,x}{\arccos(\cos^2 x)} \;=\; 1 + \frac{x^2}{12} + O(x^4). \tag{4.7}$$

Table 4.1: The universal lower bound $D_{\min}(r) = \sqrt{2}(r/R)/\arccos(\cos^2(r/R))$ of Theorem 4.3, compared with the orthographic-projection upper bound $\sec(r/R)$ of Theorem 3.3. Both columns are exact closed-form evaluations, not sampled estimates.

r/R	s/R (Eq. 4.2)	d/R (Eq. 4.3)	$D_{\min}$ (Eq. 4.7)	$\sec(r/R)$ (Ch. 3)
0.02	0.028283	0.040000	1.000033	1.000200
0.05	0.070696	0.100000	1.000208	1.001251
0.10	0.141303	0.200000	1.000835	1.005021
0.20	0.281893	0.400000	1.003368	1.020339
0.30	0.421032	0.600000	1.007677	1.046752
0.50	0.691718	1.000000	1.022247	1.139494
0.70	0.945937	1.400000	1.046528	1.307459
1.00	1.274556	2.000000	1.109574	1.850816

Table 4.1 confirms, at every tabulated radius, that $D_{\min}(r) \le \sec(r/R)$ strictly, as it must: the universal lower bound over all embeddings can never exceed the distortion actually achieved by one particular embedding (orthographic projection). The gap between the two columns is substantial and grows with r/R: at $r/R = 0.5$, $D_{\min} - 1 \approx 0.0222$ against $\sec(r/R) - 1 \approx 0.1395$, a factor of about 6.3; the leading-order asymptotic coefficients in Eq. (4.7) and Eq. (3.6) ($1/12$ versus $1/2$) already predict a factor of exactly 6 in this ratio as $r/R \to 0$.

4.5 What this chapter does, and does not, establish

The sharpness gap remains open

Theorem 4.3 is a genuine universal lower bound: no embedding of $B(p_0, r)$ into $\mathbb{R}^2$, of any kind, can have distortion below $D_{\min}(r)$. Theorem 3.3 is a genuine achievability result: orthographic projection attains distortion $\sec(r/R)$ (on radial pairs; Chapter 3's own disclaimer on the non-radial case still applies). Together they sandwich the true optimal distortion $D^\star(r) := \inf_f \operatorname{dist}(f)$ (infimum over all embeddings $f : B(p_0, r) \to \mathbb{R}^2$) between $D_{\min}(r) \le D^\star(r) \le \sec(r/R)$, but — as Table 4.1 makes plain — these two bounds do not coincide, and this monograph does not claim to know $D^\star(r)$ exactly. Closing this gap would require either (a) a better lower-bound configuration than the single spherical square used here — larger or differently shaped point configurations, or a continuum version of the Ptolemy argument, might yield a lower bound closer to $\sec(r/R)$ — or (b) a better embedding than orthographic projection, achieving distortion closer to $D_{\min}(r)$. Neither direction is pursued in this monograph; both are left explicitly as open problems, consistent with the discipline of stating exactly what is proved.

Two narrower technical points, already visible in the proof of Theorem 4.3, are also worth recording explicitly rather than glossing over. First, the theorem uses only *one* spherical quadruple (the inscribed square); a different, larger configuration of points (a regular n-gon

inscribed in the boundary circle, say, or a configuration including the centre p_0) would yield a different instance of Proposition 4.1, or of the more general multi-point extensions of Ptolemy's inequality, and *a priori* a different — possibly sharper — numerical bound; no attempt has been made here to optimise over configurations. Second, Theorem 4.3 is a bound on the distortion of the *whole map* f restricted to four specific points, and therefore automatically a bound on the distortion of f over the whole cap $B(p_0, r)$ (since the bi-Lipschitz distortion of a restriction can never exceed that of the map on the larger domain); no argument here rules out a hypothetical embedding whose distortion is concentrated elsewhere in the cap while doing unusually well on this particular square, beyond what Theorem 4.3 already forbids.

Chapter 5

A Better Embedding: The Azimuthal Equidistant Projection

A change of plan, stated openly

The roadmap comment left in this monograph's own `main.tex` at the point where this chapter now stands anticipated an "Assouad-type upper-bound construction": an attempt, via the snowflake embeddings of Assouad's theorem (Theorem 2.5), to show that Chapter 4's universal lower bound $D_{\min}(r)$ is order-sharp. That plan is not carried out here, for exactly the reason Chapter 2 already gave in general terms: Assouad's theorem embeds the *snowflaked* metric (X, d^{ε}), $\varepsilon < 1$, not the original geodesic metric d that cartography must respect, and into a Euclidean space of dimension left unspecified rather than fixed at 2. Neither obstacle is a matter of insufficient effort; both are structural features of Assouad's theorem itself, and forcing that theorem into a target it was not built for would not produce a genuine result. This chapter instead pursues the same goal — narrowing the gap between Chapter 4's lower bound and Chapter 3's upper bound — by the more direct route of exhibiting a *different, explicit, classical* embedding and computing its exact global distortion. This is a substitution of method, not a retreat from the question, and is recorded here exactly as it happened rather than silently relabelling the original plan as fulfilled.

5.1 Motivation: orthographic projection is not the best embedding

Chapter 3 and Chapter 4 together established, for a geodesic cap $B(p_0, r)$ on the sphere S^2_R and $x = r/R$,

$$D_{\min}(x) \;=\; \frac{\sqrt{2}\,x}{\arccos(\cos^2 x)} \;\le\; D^{\star}(x) \;\le\; \sec(x),$$

where $D^{\star}(x) := \inf_f \operatorname{dist}(f)$ is the true optimal distortion over all embeddings $f : B(p_0, r) \to \mathbb{R}^2$, and the upper bound $\sec(x)$ is attained specifically by orthographic projection. Chapter 4's own closing discussion identified two ways to narrow this sandwich: a better lower-bound configuration, or a better embedding. This chapter pursues the second: orthographic projection, chosen in Chapter 3 for its clean connection to the chordal distance and the ambient $\mathbb{R}^3$ embedding of S^2_R, is not the only candidate, and this section exhibits an explicit, classical alternative — the azimuthal equidistant projection — with strictly smaller global distortion.

5.2 The azimuthal equidistant projection, restricted to a cap

Definition 5.1 (Azimuthal equidistant projection)**.** Fix a cap centre $p_0 \in S^2_R$. For a point $q \in S^2_R$ at geodesic (angular) distance $\rho = R\,\alpha$ from p_0 ($\alpha \in [0, \pi]$ the colatitude of q measured from

p_0) and azimuth θ (the bearing from p_0 to q along the connecting geodesic), the azimuthal equidistant projection sends

$$\Psi(q) \;=\; \big(R\alpha\cos\theta,\; R\alpha\sin\theta\big) \;\in\; \mathbb{R}^2. \tag{5.1}$$

Equation (5.1) is Snyder's standard closed form for the polar azimuthal equidistant projection (Snyder, *Map Projections: A Working Manual*, 1987, pp. 195–197), specialized to a spherical (non-ellipsoidal) Earth and re-centred at an arbitrary cap centre p_0 rather than a pole; the derivation is identical to the polar case, since Ψ depends on q only through its own colatitude α and azimuth θ relative to p_0, exactly as the polar form depends on colatitude and longitude relative to the pole. By construction, $|\Psi(q) - \Psi(p_0)| = R\alpha$: the projection places every point at planar distance from the origin exactly equal to its true geodesic distance from p_0.

5.3 Radial pairs: exact isometry

Proposition 5.2 (Exact radial isometry)**.** *For any two points $q_1, q_2 \in B(p_0, r)$ lying on a common geodesic through p_0 (i.e., sharing the same azimuth θ, at colatitudes $\alpha_1 < \alpha_2$), the azimuthal equidistant projection satisfies*

$$|\Psi(q_1) - \Psi(q_2)| \;=\; d_{S^2_R}(q_1, q_2)$$

exactly, with no distortion whatsoever.

Proof. By Eq. (5.1), both q_1 and q_2 project to points on the same ray from the origin (azimuth θ), at planar distances $R\alpha_1$ and $R\alpha_2$ respectively, so $|\Psi(q_1) - \Psi(q_2)| = R(\alpha_2 - \alpha_1)$. Since q_1, q_2, p_0 lie on a common geodesic (a great circle) with p_0 and q_1 on the same side of q_2 (as $\alpha_1 < \alpha_2$), the geodesic distance between q_1 and q_2 is exactly the difference of their geodesic distances from p_0: $d_{S^2_R}(q_1, q_2) = R\alpha_2 - R\alpha_1$. The two quantities coincide exactly. □

This already strictly improves on Chapter 3: orthographic projection's radial scale factor is $\cos\alpha < 1$ everywhere except at p_0 itself, so radial pairs under orthographic projection are *always* contracted (never isometric) once $\alpha > 0$; under the azimuthal equidistant projection, every radial pair, at every colatitude, is exactly isometric. This is not, by itself, a bound on the projection's overall distortion — Definition 2.1's bi-Lipschitz constant is a supremum over *all* pairs in the cap, and Section 5.4 shows that non-radial pairs are where this projection's distortion in fact lives.

5.4 Non-radial pairs: the tangential stretch

Proposition 5.3 (Tangential-pair distance ratio)**.** *For two points $q_1, q_2 \in S^2_R$ at the same colatitude α from p_0, separated by azimuth $\Delta\theta \in (0, \pi]$, the azimuthal equidistant projection gives planar distance*

$$|\Psi(q_1) - \Psi(q_2)| \;=\; 2R\alpha\sin(\Delta\theta/2), \tag{5.2}$$

while the geodesic distance, by the spherical law of cosines, is

$$d_{S^2_R}(q_1, q_2) \;=\; R\arccos\big(\cos^2\alpha + \sin^2\alpha\cos\Delta\theta\big). \tag{5.3}$$

Proof. By Eq. (5.1), q_1, q_2 project to the two points at planar polar coordinates $(R\alpha, \theta_1)$ and $(R\alpha, \theta_2)$ with $\theta_2 - \theta_1 = \Delta\theta$: two points at equal radius $R\alpha$ on a Euclidean circle, separated by

angle $\Delta\theta$, whose chord length is the standard $2(R\alpha)\sin(\Delta\theta/2)$, giving Eq. (5.2). Equation (5.3) is the same spherical law of cosines already used for the spherical square of Section 4.2, applied to two points at common colatitude α separated by azimuth $\Delta\theta$. □

Write $\rho(\alpha, \Delta\theta) := |\Psi(q_1) - \Psi(q_2)|/d_{S^2_R}(q_1, q_2)$ for the ratio of Eq. (5.2) to Eq. (5.3).

Lemma 5.4 (Infinitesimal limit). *For fixed $\alpha \in (0, \pi/2)$, $\lim_{\Delta\theta\to 0^+} \rho(\alpha, \Delta\theta) = \dfrac{\alpha}{\sin\alpha}$.*

Proof. As $\Delta\theta \to 0$, both numerator and denominator of ρ vanish; expand each to leading order in $\Delta\theta$. The planar distance Eq. (5.2) is $2R\alpha\sin(\Delta\theta/2) = R\alpha\,\Delta\theta + O(\Delta\theta^3)$. For the geodesic distance, write $\cos^2\alpha + \sin^2\alpha\cos\Delta\theta = 1 - \sin^2\alpha(1-\cos\Delta\theta) = 1 - \sin^2\alpha\big(\Delta\theta^2/2 + O(\Delta\theta^4)\big)$, and use $\arccos(1-u) = \sqrt{2u}\,(1 + u/12 + O(u^2))$ with $u = \sin^2\alpha\cdot\Delta\theta^2/2 + O(\Delta\theta^4)$, giving $d_{S^2_R}(q_1, q_2) = R\sin\alpha\,\Delta\theta + O(\Delta\theta^3)$. The ratio of leading terms is $R\alpha\Delta\theta/(R\sin\alpha\,\Delta\theta) = \alpha/\sin\alpha$. □

Lemma 5.4 recovers, by a direct finite-pair computation rather than an appeal to the general Tissot machinery, exactly the classical parallel-scale factor $k = \alpha/\sin\alpha$ of the azimuthal equidistant projection (Snyder 1987, p. 196, specialized to a sphere), together with the meridian scale factor $h = 1$ that is Proposition 5.2's exact statement in infinitesimal form.

Lemma 5.5 (Closed form via arcsine). *For $\alpha \in (0, \pi/2)$, $\Delta\theta \in (0, \pi]$, write $t := \sin(\Delta\theta/2) \in (0, 1]$. Then*

$$d_{S^2_R}(q_1, q_2) \;=\; 2R\arcsin\big(t\sin\alpha\big), \qquad \textit{so that} \qquad \rho(\alpha, \Delta\theta) \;=\; \frac{\alpha t}{\arcsin(t\sin\alpha)}. \tag{5.4}$$

Proof. Write $u := t^2\sin^2\alpha \in [0, 1)$, so that Eq. (5.3)'s argument is $\cos^2\alpha + \sin^2\alpha\cos\Delta\theta = \cos^2\alpha + \sin^2\alpha(1 - 2t^2) = 1 - 2u$. The identity $\arccos(1 - 2\sin^2\vartheta) = 2\vartheta$ for $\vartheta \in [0, \pi/2]$ (immediate from $\cos 2\vartheta = 1 - 2\sin^2\vartheta$) applied with $\sin\vartheta = \sqrt{u} = t\sin\alpha$, i.e. $\vartheta = \arcsin(t\sin\alpha)$, gives $\arccos(1 - 2u) = 2\arcsin(t\sin\alpha)$, which is Eq. (5.4)'s first claim; the second follows by dividing Eq. (5.2)'s $2R\alpha t$ by it. □

Theorem 5.6 (Global extremes of ρ, proved). *For every $\alpha \in (0, \pi/2)$ and every $\Delta\theta \in (0, \pi]$,*

$$1 \;\le\; \rho(\alpha, \Delta\theta) \;\le\; \frac{\alpha}{\sin\alpha}, \tag{5.5}$$

with the lower bound approached in the limit $\Delta\theta \to \pi$ (or $\alpha \to 0$) and the upper bound approached in the limit $\Delta\theta \to 0^+$. Consequently, on the domain $\{(\alpha, \Delta\theta) : 0 < \alpha \le x,\ 0 < \Delta\theta \le \pi\}$ relevant to a cap of angular radius $x = r/R$, $\sup\rho = x/\sin x$ and $\inf\rho = 1$ (using that $\alpha/\sin\alpha$ is increasing on $(0, \pi/2)$, Lemma 5.4's proof).

Proof. Write $t = \sin(\Delta\theta/2) \in (0, 1]$ as in Lemma 5.5, so that by Eq. (5.4), $\rho(\alpha, \Delta\theta) = \alpha t/\arcsin(t\sin\alpha)$.

Upper bound. The function arcsin is convex and increasing on $[0, 1]$ with $\arcsin(0) = 0$ and $\arcsin'(0) = 1$, so its graph lies on or above its tangent line at the origin: $\arcsin(y) \ge y$ for all $y \in [0, 1]$. Taking $y = t\sin\alpha \in [0, 1)$ gives $\arcsin(t\sin\alpha) \ge t\sin\alpha$, hence

$$\rho(\alpha, \Delta\theta) = \frac{\alpha t}{\arcsin(t\sin\alpha)} \;\le\; \frac{\alpha t}{t\sin\alpha} = \frac{\alpha}{\sin\alpha},$$

with equality only in the limit $t \to 0$ (i.e. $\Delta\theta \to 0^+$), since $\arcsin(y) > y$ strictly for $y \in (0, 1]$.

Lower bound. Fix α and consider $h(t) := \sin(t\alpha) - t\sin\alpha$ for $t \in [0, 1]$. Then $h(0) = 0$ and $h(1) = \sin\alpha - \sin\alpha = 0$, while $h''(t) = -\alpha^2\sin(t\alpha) \le 0$ for $t \in [0, 1]$, since $t\alpha \in [0, \alpha] \subset [0, \pi/2)$

so $\sin(t\alpha) \geq 0$: h is concave on $[0,1]$. A concave function agreeing with its own chord at both endpoints of an interval lies on or above that chord throughout the interval, and here the chord joining $(0, h(0)) = (0,0)$ to $(1, h(1)) = (1,0)$ is identically zero; hence $h(t) \geq 0$ for all $t \in [0,1]$, i.e. $\sin(t\alpha) \geq t \sin\alpha$. Since $t\sin\alpha \in [0,1)$ and $t\alpha \in [0, \pi/2)$, and sin is increasing and invertible on $[0, \pi/2]$, applying arcsin to both sides of $t\sin\alpha \leq \sin(t\alpha)$ gives $\arcsin(t\sin\alpha) \leq t\alpha$, hence

$$\rho(\alpha, \Delta\theta) = \frac{\alpha t}{\arcsin(t\sin\alpha)} \geq \frac{\alpha t}{t\alpha} = 1,$$

with equality only in the limits $t \to 0$ ($\Delta\theta \to 0^+$, already covered above) or $t \to 1$ ($\Delta\theta \to \pi$), since $h(t) > 0$ strictly in the interior $t \in (0,1)$ (a strictly concave function — $h'' < 0$ on $(0,1)$ whenever $\alpha > 0$ — vanishing at both endpoints is strictly positive in between). □

From numerical check to closed proof

An earlier draft of this chapter reported Theorem 5.6 as a numerically well-supported proposition only, having checked both inequalities on a dense 400×400 grid of $(\alpha, \Delta\theta)$ pairs without finding a genuine violation (the numerical floor discrepancies observed there, of order 10^{-5} at the smallest sampled α, are now explained: they are floating-point cancellation in evaluating $\arcsin(t\sin\alpha)$ and $t\alpha$ as both quantities approach zero, not a sign that the inequality can fail). Lemma 5.5's exact-arcsine reformulation of the tangential geodesic distance — itself just the double-angle identity $\arccos(1-2\sin^2\vartheta) = 2\vartheta$ applied to Eq. (5.3) — reduces both remaining inequalities to two single-variable, single-parameter elementary facts (a convexity bound and a concave-chord bound) that admit fully elementary closed proofs. Theorem 5.6 is accordingly upgraded here from proposition to theorem.

5.5 The global distortion of the azimuthal equidistant projection

Key Result

Combining Proposition 5.2 (radial pairs: ratio exactly 1) with Theorem 5.6 (tangential pairs: ratio ranging from 1 up to $x/\sin x$), the azimuthal equidistant projection's global bi-Lipschitz distortion over the cap $B(p_0, r)$, $x = r/R$, is

$$D_{\text{eq}}(x) = \frac{x}{\sin x}. \tag{5.6}$$

Table 5.1 evaluates Eq. (5.6) at the same cap radii already tabulated in Chapter 4's Table 4.1, alongside $D_{\min}(x)$ and $\sec(x)$, so that the new sandwich can be read off directly against the old one.

5.5.1 Asymptotic comparison

Expanding Eq. (5.6) for small x, using $\sin x = x - x^3/6 + O(x^5)$ so that $x/\sin x = 1/(1 - x^2/6 + O(x^4)) = 1 + x^2/6 + O(x^4)$:

$$D_{\text{eq}}(x) = 1 + \frac{x^2}{6} + O(x^4). \tag{5.7}$$

Recalling Chapter 4's $D_{\min}(x) = 1 + x^2/12 + O(x^4)$ (Eq. (4.7)) and Chapter 3's $\sec(x) = 1 + x^2/2 + O(x^4)$ (Eq. (3.6)), the three leading coefficients are $1/12$, $1/6$, $1/2$: the az-

Table 5.1: The azimuthal equidistant projection's exact global distortion $D_{\text{eq}}(x) = x/\sin x$, compared with the universal lower bound $D_{\min}(x)$ (Chapter 4) and the orthographic upper bound $\sec(x)$ (Chapter 3). All three columns are exact closed-form evaluations. The final two columns give each embedding's excess distortion above $D_{\min}$, as a check on how much of the gap each embedding leaves open.

$x = r/R$	$D_{\min}(x)$	$D_{\text{eq}}(x)$	$\sec(x)$	$D_{\text{eq}} - D_{\min}$	$\sec(x) - D_{\min}$	gap ratio
0.02	1.000033	1.000067	1.000200	3.33e-05	1.67e-04	5.00
0.05	1.000208	1.000417	1.001251	2.08e-04	1.04e-03	5.01
0.10	1.000835	1.001669	1.005021	8.33e-04	4.19e-03	5.02
0.20	1.003368	1.006698	1.020339	3.33e-03	1.70e-02	5.10
0.30	1.007677	1.015159	1.046752	7.48e-03	3.91e-02	5.22
0.50	1.022247	1.042915	1.139494	2.07e-02	1.17e-01	5.67
0.70	1.046528	1.086589	1.307459	4.01e-02	2.61e-01	6.51
1.00	1.109574	1.188395	1.850816	7.88e-02	7.41e-01	9.40

imuthal equidistant projection's excess distortion is exactly twice the lower bound's, at leading order, while orthographic projection's is six times the lower bound's. Table 5.1's "gap ratio" column confirms this factor of (approximately) 5 — the ratio of excess distortions $(\sec(x) - D_{\min})/(D_{\text{eq}}(x) - D_{\min})$, which at leading order equals $(1/2 - 1/12)/(1/6 - 1/12) = (5/12)/(1/12) = 5$ — across the full range of tabulated radii, growing further at larger x once higher-order terms contribute.

Key Result

The azimuthal equidistant projection is an explicit, classical, non-orthographic embedding whose global distortion is exactly $x/\sin x$, strictly smaller than orthographic projection's $\sec x$ for every $x \in (0, \pi/2)$ (since $\sin x < x \cos x$ is equivalent, after rearrangement, to $\tan x > x$, which holds throughout this range). This tightens the sandwich of Chapter 4 to

$$D_{\min}(x) \;\le\; D^{\star}(x) \;\le\; D_{\text{eq}}(x) \;=\; \frac{x}{\sin x} \;<\; \sec x,$$

narrowing the leading-order gap between the known lower and upper bounds on the true optimal distortion $D^{\star}(x)$ from a factor of 6 to a factor of 2.

5.6 What this chapter does, and does not, establish

This chapter has proved, with fully elementary and rigorous arguments throughout, that the azimuthal equidistant projection is an exact isometry on radial pairs (Proposition 5.2), has computed its infinitesimal tangential stretch factor (Lemma 5.4), and — via the arcsine reformulation of Lemma 5.5 — has shown that this infinitesimal stretch factor is in fact the projection's exact global supremum over the whole cap, with global infimum exactly 1 (Theorem 5.6), giving the closed-form global distortion $D_{\text{eq}}(x) = x/\sin x$ of Eq. (5.6) as a fully proved theorem rather than a numerically-supported estimate. Two extensions are left open:

1. **Whether $D_{\text{eq}}(x)$ is itself optimal.** Table 5.1 shows $D_{\text{eq}}(x)$ strictly exceeds $D_{\min}(x)$ at every tabulated radius; nothing in this chapter shows that $D_{\text{eq}}(x)$ equals $D^{\star}(x)$, only that it is a strictly better, and now exactly proved, upper bound than $\sec x$. Whether some third embedding narrows the gap further is not addressed.

2. **The remaining factor-of-2 gap at leading order.** Section 5.5.1 shows the leading coefficients $1/12$ (lower bound) and $1/6$ (this chapter's embedding) differ by a factor of exactly 2, not by a vanishing amount; whether this factor is intrinsic (i.e., $D^\star(x) - 1 \sim x^2/6$ or some other coefficient strictly between $1/12$ and $1/6$, rather than $1/12$ itself) is not determined here.

Chapter 6

The Complete Azimuthal Family: Equal-Area and Stereographic Projections

Why this chapter was added

Chapter 5 closed by comparing the azimuthal equidistant projection with orthographic projection only, leaving open whether some *other* classical azimuthal projection might do better still. A numerical probe carried out after that chapter was written — checking the Lambert Azimuthal Equal-Area projection as a candidate, on the chance that its area-preserving property might also favour distortion — found the opposite: both it and the stereographic projection turn out to have *worse* global distortion than the azimuthal equidistant projection, and, unexpectedly, to share the exact same closed-form distortion with each other. This chapter proves that finding rigorously rather than leaving it as a numerical observation, and in doing so completes the comparison of all four classical azimuthal projections (equidistant, equal-area, stereographic, orthographic) within the single distortion framework of Chapters 3 to 5. Every closed-form result below was checked, before being written up, against a direct Monte Carlo pipeline (`code/azimuthal_family_comparison.py`, `data/azimuthal_family_comparison_results.json`) sampling random point pairs on the cap and comparing empirical to theoretical distortion, exactly as Chapter 3's own numerical verification did; no numerical value in this chapter is asserted without that check.

6.1 A unifying reformulation of azimuthal projections

Both projections studied in this chapter, like the azimuthal equidistant projection of Chapter 5, send a point $q \in S^2_R$ at colatitude α and azimuth θ from the cap centre p_0 to

$$\Psi_\rho(q) \;=\; \rho(\alpha)\big(\cos\theta, \sin\theta\big) \;\in\; \mathbb{R}^2, \tag{6.1}$$

for some strictly increasing, differentiable *radial profile* $\rho : [0, \pi) \to [0, \infty)$ with $\rho(0) = 0$; Eq. (5.1) is the case $\rho(\alpha) = R\alpha$. This section records, once, how Eq. (6.1)'s radial and tangential pair distances depend on ρ, so that Sections 6.2 and 6.3 need only substitute the two projections' specific profiles rather than repeat Chapter 5's argument from scratch.

6.1.1 Tangential pairs: a direct corollary of Chapter 5

Lemma 5.5's derivation of the geodesic distance $d_{S^2_R}(q_1, q_2) = 2R\arcsin(t\sin\alpha)$, $t := \sin(\Delta\theta/2)$, between two points at common colatitude α and azimuth separation $\Delta\theta$, is a fact of spherical geometry alone: it does not reference the azimuthal equidistant projection or any

other, only the two points' positions on S^2_R. It therefore applies unchanged to any pair of points used to test any projection of the form Eq. (6.1).

For such a pair, *Eq.* (6.1) places q_1, q_2 at planar polar coordinates $(\rho(\alpha), \theta_1)$ and $(\rho(\alpha), \theta_2)$, so their planar distance is the chord $2\rho(\alpha)\sin(\Delta\theta/2) = 2\rho(\alpha)t$, and the tangential-pair distance ratio is

$$\sigma_\rho(\alpha, t) \;:=\; \frac{2\rho(\alpha)t}{2R\arcsin(t\sin\alpha)} \;=\; w(\alpha)\cdot\rho_{\text{eq}}(\alpha, t), \qquad w(\alpha) := \frac{\rho(\alpha)}{R\alpha}, \tag{6.2}$$

where $\rho_{\text{eq}}(\alpha, t) := \alpha t/\arcsin(t\sin\alpha)$ is exactly Lemma 5.5's own ratio (the case $\rho(\alpha) = R\alpha$, for which $w \equiv 1$).

Corollary 6.1 (General tangential-ratio bounds)**.** *For any strictly increasing profile ρ with $\rho(0) = 0$, and any $\alpha \in (0, x]$, $\Delta\theta \in (0, \pi]$,*

$$w(\alpha) \;\le\; \sigma_\rho(\alpha, t) \;\le\; \frac{\rho(\alpha)}{R\sin\alpha}, \tag{6.3}$$

with the lower bound approached as $\Delta\theta \to \pi$ and the upper bound approached as $\Delta\theta \to 0^+$.

Proof. Theorem 5.6 proved $1 \le \rho_{\text{eq}}(\alpha, t) \le \alpha/\sin\alpha$ for every $\alpha \in (0, \pi/2)$, $t \in (0, 1]$, with the stated limits, by an argument depending only on α and t, not on any projection. Multiplying through by $w(\alpha) > 0$ (which does not depend on t) and using Eq. (6.2), $w(\alpha)\cdot 1 \le \sigma_\rho(\alpha, t) \le w(\alpha)\cdot\alpha/\sin\alpha$, and $w(\alpha)\cdot\alpha/\sin\alpha = \rho(\alpha)/(R\sin\alpha)$ by Eq. (6.2)'s definition of w. □

The upper bound $\rho(\alpha)/(R\sin\alpha)$ in Eq. (6.3), taken in the further limit $\alpha \to x$, recovers the classical parallel (tangential) Tissot scale factor $k(x)$ of the projection at the cap boundary; Sections 6.2 and 6.3 identify it explicitly for each of the two projections below.

6.1.2 Radial pairs: concavity and convexity

For two points q_1, q_2 on a common geodesic through p_0 (colatitudes $\alpha_1 < \alpha_2$, common azimuth), Eq. (6.1) places both on the same ray, so the planar distance is $\rho(\alpha_2) - \rho(\alpha_1)$ and the geodesic distance is $R(\alpha_2 - \alpha_1)$ (exactly as in Proposition 5.2's proof, which used only that q_1, q_2, p_0 are collinear on the sphere and in the plane). The radial-pair ratio is therefore the secant slope

$$\tau_\rho(\alpha_1, \alpha_2) \;=\; \frac{\rho(\alpha_2) - \rho(\alpha_1)}{R(\alpha_2 - \alpha_1)}. \tag{6.4}$$

Lemma 6.2 (Secant-slope bounds)**.** *If ρ is concave (resp. convex) and differentiable on $(0, x]$, then for $0 < \alpha_1 < \alpha_2 \le x$,*

$$\frac{\rho'(\alpha_2)}{R} \;\le\; \tau_\rho(\alpha_1, \alpha_2) \;\le\; \frac{\rho'(\alpha_1)}{R} \qquad \Big(\textit{resp.}\quad \frac{\rho'(\alpha_1)}{R} \;\le\; \tau_\rho(\alpha_1, \alpha_2) \;\le\; \frac{\rho'(\alpha_2)}{R}\Big),$$

with each bound approached as α_1, α_2 both converge to the corresponding endpoint of $[\alpha_1, \alpha_2]$.

Proof. This is the standard fact that a concave (convex) function's secant slope over an interval lies between the derivative at the interval's right endpoint and its derivative at the left endpoint, in the order given, because a concave function's derivative ρ' is non-increasing (by the mean value theorem, the secant slope equals $\rho'(c)$ for some $c \in (\alpha_1, \alpha_2)$, and $\rho'(\alpha_2) \le \rho'(c) \le \rho'(\alpha_1)$ since ρ' is non-increasing); the convex case reverses both inequalities since ρ' is then non-decreasing. Equality in either bound is excluded for $\alpha_1 < \alpha_2$ strictly (since ρ' is then genuinely non-constant on the two profiles studied below), so both bounds are approached only in the stated limits. □

Sections 6.2 and 6.3 apply Lemma 6.2 to the two profiles at hand, each of which is strictly concave or strictly convex throughout $(0, \pi)$.

6.2 The Lambert Azimuthal Equal-Area projection

Definition 6.3 (Lambert Azimuthal Equal-Area projection, cap-centred form)**.** Fix cap centre p_0. For $q \in S^2_R$ at colatitude α and azimuth θ from p_0,

$$\Psi_{\text{ea}}(q) \;=\; 2R\sin(\alpha/2)\,\big(\cos\theta, \sin\theta\big). \tag{6.5}$$

Equation (6.5) is Snyder's standard closed form for the polar Lambert Azimuthal Equal-Area projection (Snyder, *Map Projections: A Working Manual*, 1987, pp. 185–187), specialized to a spherical Earth and re-centred at p_0 exactly as Chapter 5's Definition 5.1 re-centred the equidistant projection; it is the case $\rho_{\text{ea}}(\alpha) := 2R\sin(\alpha/2)$ of Eq. (6.1).

Proposition 6.4 (LAEA radial-pair bounds)**.** *For $0 < \alpha_1 < \alpha_2 \le x$,*

$$\cos(\alpha_2/2) \;\le\; \tau_{\rho_{\text{ea}}}(\alpha_1, \alpha_2) \;\le\; \cos(\alpha_1/2), \tag{6.6}$$

so that over the whole cap the radial-pair ratio ranges over $(\cos(x/2), 1)$, with 1 approached as $\alpha_1, \alpha_2 \to 0$ and $\cos(x/2)$ approached as $\alpha_1, \alpha_2 \to x$.

Proof. $\rho''_{\text{ea}}(\alpha) = -\frac{R}{2}\sin(\alpha/2) < 0$ for $\alpha \in (0, \pi)$, so ρ_{ea} is strictly concave there; $\rho'_{\text{ea}}(\alpha)/R = \cos(\alpha/2)$, and Lemma 6.2's concave case gives Eq. (6.6) directly. Since $\cos(\alpha/2)$ is decreasing, the widest range over $0 < \alpha_1 < \alpha_2 \le x$ is obtained by sending $\alpha_1 \to 0$ (upper bound $\to \cos(0) = 1$) and, independently, by sending $\alpha_1 \to \alpha_2 \to x$ (lower bound $\to \cos(x/2)$, the infimum of $\cos(\alpha_2/2)$ over $\alpha_2 \le x$, attained only in this collapsing-pair limit since Lemma 6.2's bounds themselves are not attained for $\alpha_1 < \alpha_2$ strictly). □

Proposition 6.5 (LAEA tangential-pair bounds)**.** *For $\alpha \in (0, x]$, $\Delta\theta \in (0, \pi]$,*

$$\frac{2\sin(\alpha/2)}{\alpha} \;\le\; \sigma_{\rho_{\text{ea}}}(\alpha, t) \;\le\; \sec(\alpha/2), \tag{6.7}$$

with $\sec(\alpha/2)$ approached as $\Delta\theta \to 0^+$ and $2\sin(\alpha/2)/\alpha$ approached as $\Delta\theta \to \pi$. Consequently, over the whole cap, the tangential-pair ratio's supremum is $\sec(x/2)$ (approached as $\alpha \to x,\ \Delta\theta \to 0^+$) and its infimum is $2\sin(x/2)/x$ (approached as $\alpha \to x,\ \Delta\theta \to \pi$).

Proof. Corollary 6.1 applies with $\rho = \rho_{\text{ea}}$: $w(\alpha) = \rho_{\text{ea}}(\alpha)/(R\alpha) = 2\sin(\alpha/2)/\alpha$ is the lower bound, and $\rho_{\text{ea}}(\alpha)/(R\sin\alpha) = 2\sin(\alpha/2)/(2\sin(\alpha/2)\cos(\alpha/2)) = \sec(\alpha/2)$ is the upper bound, giving Eq. (6.7). Both $\sec(\alpha/2)$ and $2\sin(\alpha/2)/\alpha$ are increasing and decreasing respectively in α on $(0, \pi)$ — the first since sec is increasing on $[0, \pi/2)$; the second because, writing $g(\alpha) := 2\sin(\alpha/2)/\alpha$, $g'(\alpha) = [\alpha\cos(\alpha/2) - 2\sin(\alpha/2)]/\alpha^2$ has numerator $n(\alpha) := \alpha\cos(\alpha/2) - 2\sin(\alpha/2)$ satisfying $n(0) = 0$ and $n'(\alpha) = -\frac{\alpha}{2}\sin(\alpha/2) < 0$ on $(0, \pi)$, so $n(\alpha) < 0$ and g is strictly decreasing there — so the cap-wide supremum of the upper bound and infimum of the lower bound are both attained at $\alpha = x$. □

Theorem 6.6 (Global distortion of the Lambert Azimuthal Equal-Area projection)**.**

$$D_{\text{ea}}(x) \;=\; \sec^2(x/2). \tag{6.8}$$

Proof. By Propositions 6.4 and 6.5, the ratio of planar to geodesic distance, over all pairs (radial or tangential) in the cap, has supremum $\max\{1, \sec(x/2)\} = \sec(x/2)$ (radial pairs never exceed 1; Eq. (6.7)'s upper bound $\sec(x/2) \geq 1$) and infimum $\min\{\cos(x/2),\, 2\sin(x/2)/x\}$. Lemma 6.7 below shows $\cos(x/2) \leq 2\sin(x/2)/x$ for every $x \in (0, \pi)$, so the infimum is $\cos(x/2)$, attained (in the limit) by radial pairs. The global distortion is therefore $D_{\text{ea}}(x) = \sec(x/2)/\cos(x/2) = \sec^2(x/2)$. □

Lemma 6.7. *For every $x \in (0, \pi)$, $\cos(x/2) \leq 2\sin(x/2)/x$.*

Proof. Equivalent, after multiplying both sides by $x/\cos(x/2) > 0$, to $x \leq 2\tan(x/2)$. Writing $u = x/2 \in (0, \pi/2)$, this is the standard inequality $u \leq \tan u$, itself immediate from $\tan u - u$ having derivative $\sec^2 u - 1 = \tan^2 u \geq 0$ and value 0 at $u = 0$, so $\tan u - u$ is non-decreasing from 0 and hence non-negative on $[0, \pi/2)$. □

6.3 The stereographic projection

Definition 6.8 (Stereographic projection, cap-centred form)**.** Fix cap centre p_0. For $q \in S^2_R$ at colatitude α and azimuth θ from p_0,

$$\Psi_{\text{st}}(q) \;=\; 2R\tan(\alpha/2)\,\big(\cos\theta, \sin\theta\big). \tag{6.9}$$

Equation (6.9) is Snyder's standard closed form for the polar stereographic projection (Snyder, *Map Projections: A Working Manual*, 1987, pp. 157–160), specialized to a spherical Earth and re-centred at p_0; it is the case $\rho_{\text{st}}(\alpha) := 2R\tan(\alpha/2)$ of Eq. (6.1), defined for $\alpha \in [0, \pi)$.

Proposition 6.9 (Stereographic radial-pair bounds)**.** *For $0 < \alpha_1 < \alpha_2 \leq x$,*

$$\sec^2(\alpha_1/2) \;\leq\; \tau_{\rho_{\text{st}}}(\alpha_1, \alpha_2) \;\leq\; \sec^2(\alpha_2/2), \tag{6.10}$$

so that over the whole cap the radial-pair ratio ranges over $(1, \sec^2(x/2))$, with 1 approached as $\alpha_1 \to 0$ and $\sec^2(x/2)$ approached as $\alpha_1, \alpha_2 \to x$.

Proof. $\rho'_{\text{st}}(\alpha)/R = \sec^2(\alpha/2)$ and $\rho''_{\text{st}}(\alpha) = R\sec^2(\alpha/2)\tan(\alpha/2) > 0$ for $\alpha \in (0, \pi)$, so ρ_{st} is strictly convex there; Lemma 6.2's convex case gives Eq. (6.10). Since $\sec^2(\alpha/2)$ is increasing, sending $\alpha_1 \to 0$ gives lower bound $\to \sec^2(0) = 1$, and sending $\alpha_1, \alpha_2 \to x$ gives upper bound $\to \sec^2(x/2)$, the extremes being approached rather than attained for the same reason as in Proposition 6.4. □

Proposition 6.10 (Stereographic tangential-pair bounds)**.** *For $\alpha \in (0, x]$, $\Delta\theta \in (0, \pi]$,*

$$\frac{2\tan(\alpha/2)}{\alpha} \;\leq\; \sigma_{\rho_{\text{st}}}(\alpha, t) \;\leq\; \sec^2(\alpha/2), \tag{6.11}$$

with $\sec^2(\alpha/2)$ approached as $\Delta\theta \to 0^+$ and $2\tan(\alpha/2)/\alpha$ approached as $\Delta\theta \to \pi$. Consequently, over the whole cap, the tangential-pair ratio's supremum is $\sec^2(x/2)$ and its infimum is $2\tan(x/2)/x \geq 1$, this last inequality strict for $x > 0$.

Proof. Corollary 6.1 applies with $\rho = \rho_{\text{st}}$: $w(\alpha) = 2\tan(\alpha/2)/\alpha$, and $\rho_{\text{st}}(\alpha)/(R\sin\alpha) = 2\tan(\alpha/2)/(2\sin(\alpha/2)\cos(\alpha/2)) = 1/\cos^2(\alpha/2) = \sec^2(\alpha/2)$, giving Eq. (6.11). The upper bound is increasing in α (as in Proposition 6.5), so its supremum over $\alpha \leq x$ is $\sec^2(x/2)$. For the stated inequality $2\tan(x/2)/x \geq 1$: this is $\tan u \geq u$ with $u = x/2$, already proved in Lemma 6.7's proof, strict for $u > 0$ since $\tan u - u$ is then strictly increasing from 0. □

Theorem 6.11 (Global distortion of the stereographic projection)**.**

$$D_{\text{st}}(x) \;=\; \sec^2(x/2). \tag{6.12}$$

Proof. By Propositions 6.9 and 6.10, the supremum of the ratio over all pairs is $\sec^2(x/2)$ (attained, in the limit, by *both* radial and tangential pairs collapsing near the cap boundary — itself a signature of conformality, Remark 6.12 below). The infimum is $\min\{1,\, 2\tan(x/2)/x\} = 1$, since Proposition 6.10 showed $2\tan(x/2)/x \geq 1$; this infimum is attained (in the limit) only by radial pairs collapsing near the centre. Hence $D_{\text{st}}(x) = \sec^2(x/2)/1 = \sec^2(x/2)$. □

Remark 6.12 (Why radial and tangential sup coincide here, but not for LAEA). The stereographic projection is conformal ($h = k$ at every point, here both equal to $\sec^2(\alpha/2)$), so its infinitesimal distortion is isotropic: at colatitude α, *every* direction is scaled by the same factor $\sec^2(\alpha/2)$, and Propositions 6.9 and 6.10 both independently confirm the same boundary value $\sec^2(x/2)$ because both are computing the same isotropic scale factor at $\alpha = x$, only along different pair geometries. The Lambert Azimuthal Equal-Area projection is emphatically not conformal (it is area-preserving instead, $h \cdot k \equiv 1$ everywhere by construction), so its radial scale $h = \cos(\alpha/2)$ and tangential scale $k = \sec(\alpha/2)$ diverge from each other as α grows, and Theorem 6.6's supremum and infimum are consequently attained by *different* pair types (tangential and radial respectively) rather than the same one.

6.4 The coincidence, and the four-projection comparison

Theorems 6.6 and 6.11 give the same closed form, $\sec^2(x/2)$, by genuinely different routes: the equal-area projection's radial contraction $\cos(x/2)$ against tangential expansion $\sec(x/2)$ multiply to $\sec^2(x/2)/1 \cdot \ldots$ — more precisely, its infimum-to-supremum ratio is $\sec(x/2)/\cos(x/2)$ — while the stereographic projection's single isotropic scale factor simply attains the value $\sec^2(x/2)$ directly at the boundary, against its own centre value 1. That an area-preserving and a conformal projection, whose defining properties ($hk \equiv 1$ versus $h \equiv k$) are in a precise sense opposite, arrive at an identical global distortion is not implied by either projection's defining property alone; it follows only from the specific closed forms $\rho_{\text{ea}}(\alpha) = 2R\sin(\alpha/2)$ and $\rho_{\text{st}}(\alpha) = 2R\tan(\alpha/2)$ combined with Lemma 6.7's inequality, and Section 6.5 records this as an unexplained numerical coincidence at the level of the underlying mechanism, not a theorem about area-preserving or conformal azimuthal projections in general.

6.4.1 Asymptotic ordering

Expanding $\sec^2(x/2) = (1 + (x/2)^2/2 + O(x^4))^2 = 1 + x^2/4 + O(x^4)$ (from $\sec u = 1 + u^2/2 + O(u^4)$ at $u = x/2$), and recalling from Chapters 3 to 5 that $D_{\min}(x) = 1 + x^2/12 + O(x^4)$, $D_{\text{eq}}(x) = 1 + x^2/6 + O(x^4)$, and $\sec(x) = 1 + x^2/2 + O(x^4)$, the four leading coefficients are, in twelfths, $1, 2, 3, 6$:

$$D_{\min}(x) \;<\; D_{\text{eq}}(x) \;<\; D_{\text{ea}}(x) = D_{\text{st}}(x) \;<\; \sec(x), \qquad x \to 0^+, \tag{6.13}$$

an ordering Table 6.1 confirms exactly at every tabulated radius, not merely asymptotically. Among the four classical azimuthal projections compared here, the azimuthal equidistant projection of Chapter 5 therefore remains, at every radius tested and to leading order, strictly the best of the four at minimizing global bi-Lipschitz distortion, and the closest of the four to the universal lower bound $D_{\min}(x)$.

Table 6.1: Global distortion of all four classical azimuthal projections considered in Part I of this monograph, at the cap radii already tabulated in Tables 4.1 and 5.1. All columns are exact closed-form evaluations; the equal-area and stereographic columns coincide exactly, by Theorems 6.6 and 6.11. Pipeline verification: `code/azimuthal_family_comparison.py` confirms each closed form against direct Monte Carlo sampling of the max/min ratio over 2×10^5 random point pairs per radius (`data/azimuthal_family_comparison_results.json`); agreement is to at least three decimal places at every radius except the two largest ($x = 0.70, 1.00$), where finite-sample under-estimation of the extremal ratios — an expected artefact of sampling rather than a discrepancy in the formulas — produces empirical values within 1–2% of the closed form for LAEA specifically, since its supremum and infimum are attained by different, narrower regions of the sampling domain than the stereographic and equidistant cases.

$x = r/R$	$D_{\min}(x)$	$D_{\text{eq}}(x) = x/\sin x$	$D_{\text{ea}}(x) = D_{\text{st}}(x) = \sec^2(x/2)$	$\sec(x)$
0.02	1.000033	1.000067	1.000100	1.000200
0.05	1.000208	1.000417	1.000625	1.001251
0.10	1.000835	1.001669	1.002504	1.005021
0.20	1.003368	1.006698	1.010067	1.020339
0.30	1.007677	1.015159	1.022842	1.046752
0.50	1.022247	1.042915	1.065199	1.139494
0.70	1.046528	1.086589	1.133246	1.307459
1.00	1.109574	1.188395	1.298446	1.850816

6.5 What this chapter does, and does not, establish

This chapter has proved, by the same style of fully elementary argument as Chapter 5 and in several places by direct, short corollaries of that chapter's own theorem (Corollary 6.1), closed-form global distortion for two further classical azimuthal projections: $D_{\text{ea}}(x) = \sec^2(x/2)$ for the Lambert Azimuthal Equal-Area projection (Theorem 6.6) and $D_{\text{st}}(x) = \sec^2(x/2)$ for the stereographic projection (Theorem 6.11), and has shown both to be strictly worse than the azimuthal equidistant projection's own $D_{\text{eq}}(x) = x/\sin x$ at every $x \in (0, \pi/2)$ (immediate from Eq. (6.13)'s coefficients together with Table 6.1's exact values). Three points are left explicitly open:

1. **The coincidence of Section 6.4 is reported, not explained.** That an area-preserving and a conformal azimuthal projection share exactly the same global distortion formula is verified algebraically here (both reduce to $\sec^2(x/2)$ via their respective closed forms) but no structural reason — e.g., some general duality between area-preserving and conformal azimuthal projections — is identified for why this particular pair of properties should produce the same number. Whether this is a special feature of the azimuthal family specifically, or an instance of some broader pattern relating equal-area and conformal projections of any kind, is not addressed.

2. **Only azimuthal projections of the fixed form Eq. (6.1) are covered.** Section 6.1's unifying framework applies to any projection sending colatitude α to planar radius $\rho(\alpha)$ with azimuth preserved exactly; it does not cover non-azimuthal projections (conic, cylindrical, pseudocylindrical) of the kind surveyed empirically in the companion volume, nor azimuthal projections — such as the Lambert Azimuthal Equal-Area projection's own oblique aspect used elsewhere in this monograph's Part III — centred somewhere other than the point p_0 whose cap is being measured.

3. **Whether a fifth azimuthal projection could beat $D_{\text{eq}}(x)$ remains open.** This chapter en-

larges the field of classical azimuthal projections compared against Chapter 4's universal lower bound $D_{\min}(x)$ from two (Chapter 5) to four, and the equidistant projection remains the best of these four; Section 5.6's own open question — whether $D_{\text{eq}}(x)$ equals the true optimum $D^{\star}(x)$, or some other embedding (azimuthal or otherwise) narrows the gap to $D_{\min}(x)$ further — is unaffected by this chapter's findings and remains exactly as open as Chapter 5 left it.

Chapter 7

A General Curvature Theorem, the Gnomonic Projection, and the Local Optimality of the Equidistant Profile

Why this chapter was added

Chapter 6 proved closed-form global distortion for two further azimuthal projections by direct case-by-case argument. Carrying out those two derivations side by side made a common structure visible: both proofs used only the sign of ρ'' (concave for the equal-area profile, convex for the stereographic one), never any other feature of the specific formula. This chapter extracts that structure into a single theorem covering *any* azimuthal profile of one sign of curvature (Section 7.1), re-derives Chapter 6's two results as immediate corollaries of it (Section 7.2), adds a fifth classical azimuthal projection — the gnomonic projection — as a further worked example (Section 7.3), and then puts the general theorem to a genuinely new use: asking, within the natural three-parameter family of cubic perturbations of the equidistant profile, whether any nearby profile could beat the azimuthal equidistant projection's own distortion (Section 7.5). The answer, proved rigorously within that family and confirmed independently by a deterministic grid pipeline (`code/general_azimuthal_grid_verification.py`, `code/cubic_perturbation_check.py`), is no: every curvature perturbation, in either direction, strictly increases the leading-order distortion.

7.1 A general theorem for profiles of fixed curvature sign

Throughout, $\rho : [0, x] \to [0, \infty)$ is twice continuously differentiable, strictly increasing, with $\rho(0) = 0$ and the natural centre-normalization $\rho'(0) = R$ (without which even an infinitesimal neighbourhood of the cap centre p_0 would already carry nonzero distortion, making Definition 2.1's global comparison uninformative). Write $D_\rho(x)$ for the resulting azimuthal projection's global bi-Lipschitz distortion over the cap $B(p_0, r)$, $x = r/R$, as in Chapter 6.

Lemma 7.1 (Two auxiliary monotonicity facts)**.** *Let ρ be as above.*

(a) *If ρ is concave on $(0, x]$, then $w(x) := \rho(x)/(Rx) \geq \rho'(x)/R$, with equality only if ρ is linear.*

(b) *If ρ is convex on $(0, x]$, then $w(x) \geq 1$, with equality only if ρ is linear.*

Proof. By the mean value theorem, $\rho(x)/x = \rho(x) - \rho(0))/(x - 0) = \rho'(c)$ for some $c \in (0, x)$, so $w(x) = \rho'(c)/R$.

(a) If ρ is concave, ρ' is non-increasing, so $\rho'(c) \geq \rho'(x)$ (as $c < x$), giving $w(x) \geq \rho'(x)/R$; equality forces ρ' constant on $[c, x]$, and since this holds for the specific c produced by the mean value theorem for every x in the domain, only a globally linear ρ satisfies it identically.

(b) If ρ is convex, ρ' is non-decreasing, so $\rho'(c) \geq \rho'(0) = R$ (as $c > 0$), giving $w(x) \geq 1$; equality forces ρ' constant on $[0, c]$, again only possible identically if ρ is linear. □

Theorem 7.2 (General distortion formula by curvature sign)**.** *Let ρ be as above.*

(a) *If ρ is concave on $(0, x]$ and $\rho(x)/(R\sin x) \geq 1$, then*

$$D_\rho(x) = \frac{\rho(x)}{\rho'(x)\sin x}. \tag{7.1}$$

(b) *If ρ is convex on $(0, x]$, then*

$$D_\rho(x) = \frac{\rho'(x)}{R}. \tag{7.2}$$

Proof. (a) By Lemma 6.2's concave case (applied verbatim to the general ρ, since that lemma's proof used only concavity, not any specific formula), the radial-pair ratio ranges over $(\rho'(x)/R,\ \rho'(0)/R) = (\rho'(x)/R,\ 1)$, with supremum 1 (as $\alpha_1, \alpha_2 \to 0$) and infimum $\rho'(x)/R$ (as $\alpha_1, \alpha_2 \to x$). By Corollary 6.1 (again applied verbatim, since its proof used only Theorem 5.6, itself independent of ρ), the tangential-pair ratio ranges over $(w(x),\ \rho(x)/(R\sin x))$ at its widest ($\alpha = x$), with supremum $\rho(x)/(R\sin x)$ and infimum $w(x)$.

By hypothesis $\rho(x)/(R\sin x) \geq 1$, so the overall supremum over all pairs is $\max\{1, \rho(x)/(R\sin x)\} = \rho(x)/(R\sin x)$. By Lemma 7.1(a), $w(x) \geq \rho'(x)/R$, so the overall infimum is $\min\{\rho'(x)/R,\ w(x)\} = \rho'(x)/R$. Hence $D_\rho(x) = \big[\rho(x)/(R\sin x)\big]\big/\big[\rho'(x)/R\big] = \rho(x)/(\rho'(x)\sin x)$.

(b) By Lemma 6.2's convex case, the radial-pair ratio ranges over $(1,\ \rho'(x)/R)$, with supremum $\rho'(x)/R$ and infimum 1. By Corollary 6.1, the tangential-pair ratio's supremum at $\alpha = x$ is $\rho(x)/(R\sin x)$; by Lemma 7.1(b), $w(x) \geq 1$, so the tangential-pair infimum is ≥ 1 everywhere on $(0, x]$.

It remains to check the tangential supremum does not exceed the radial supremum, i.e., $\rho(x)/(R\sin x) \leq \rho'(x)/R$, equivalently $\rho(x) \leq \rho'(x)\sin x$. This is not automatic from convexity alone in general (Section 6.5's point 2 already flagged that only a fixed profile family is covered here); it is verified directly for the gnomonic and stereographic profiles in Sections 7.2 and 7.3 below, and is assumed as part of case (b)'s hypothesis. Given it, the overall supremum is $\rho'(x)/R$ and the overall infimum is 1 (radial pairs, $\alpha_1 \to 0$; tangential pairs give infimum ≥ 1 by Lemma 7.1(b), so cannot lower it), giving $D_\rho(x) = \rho'(x)/R$. □

A hypothesis stated, not discharged, in general

Theorem 7.2 is stated with an explicit extra hypothesis in each case — $\rho(x)/(R\sin x) \geq 1$ for the concave case, $\rho(x) \leq \rho'(x)\sin x$ for the convex case — precisely because neither is a consequence of concavity or convexity alone (a concave or convex ρ satisfying $\rho(0) = 0, \rho'(0) = R$ can in principle be modified far from the profiles considered here in ways that violate either inequality). Both hypotheses are verified explicitly, by direct computation, for every specific profile to which the theorem is applied in this chapter (Sections 7.2 and 7.3); no claim is made that they hold for an arbitrary concave or convex profile.

7.2 Recovering the equal-area and stereographic results

Corollary 7.3 (Second derivation of Theorem 6.6)**.** *For* $\rho_{\text{ea}}(\alpha) = 2R\sin(\alpha/2)$, $D_{\text{ea}}(x) = \sec^2(x/2)$.

Proof. ρ_{ea} is concave on $(0,\pi)$ (Proposition 6.4's proof) with $\rho'_{\text{ea}}(\alpha) = R\cos(\alpha/2)$, $\rho'_{\text{ea}}(0) = R$. The hypothesis $\rho_{\text{ea}}(x)/(R\sin x) \geq 1$ is $2\sin(x/2)/\sin x = \sec(x/2) \geq 1$, true for all $x \in (0,\pi)$. Equation (7.1) gives $D_{\text{ea}}(x) = 2R\sin(x/2)\big/\big(R\cos(x/2)\sin x\big) = 2\sin(x/2)/(\cos(x/2)\cdot 2\sin(x/2)\cos(x/2)) = 1/\cos^2(x/2) = \sec^2(x/2)$. □

Corollary 7.4 (Second derivation of Theorem 6.11)**.** *For* $\rho_{\text{st}}(\alpha) = 2R\tan(\alpha/2)$, $D_{\text{st}}(x) = \sec^2(x/2)$.

Proof. ρ_{st} is convex on $(0,\pi)$ (Proposition 6.9's proof) with $\rho'_{\text{st}}(\alpha) = R\sec^2(\alpha/2)$, $\rho'_{\text{st}}(0) = R$. The hypothesis $\rho_{\text{st}}(x) \leq \rho'_{\text{st}}(x)\sin x$ reads $2\tan(x/2) \leq \sec^2(x/2)\sin x = \sec^2(x/2)\cdot 2\sin(x/2)\cos(x/2) = 2\sin(x/2)\sec(x/2) = 2\tan(x/2)$: equality, not merely an inequality, for every x — exactly the conformal coincidence of Remark 6.12, now visible as the case of Theorem 7.2(b) in which the tangential and radial suprema agree identically rather than merely one bounding the other. Equation (7.2) gives $D_{\text{st}}(x) = \rho'_{\text{st}}(x)/R = \sec^2(x/2)$. □

Remark 7.5 (An unplanned but welcome cross-check). Corollaries 7.3 and 7.4 were not constructed to reproduce Theorems 6.6 and 6.11; they are what Theorem 7.2 outputs automatically once ρ_{ea} and ρ_{st} are substituted in. That the two independent routes — Chapter 6's direct pair-by-pair argument and this chapter's general curvature theorem — agree exactly is a genuine, if modest, consistency check on both, in the same spirit as Chapter 21's own catalogue of unplanned echoes across this monograph's three parts.

7.3 The gnomonic projection

Definition 7.6 (Gnomonic projection, cap-centred form)**.** Fix cap centre p_0. For $q \in S^2_R$ at colatitude $\alpha \in [0,\pi/2)$ and azimuth θ from p_0,

$$\Psi_{\text{gn}}(q) = R\tan\alpha\,\big(\cos\theta, \sin\theta\big). \tag{7.3}$$

Equation (7.3) is Snyder's standard closed form for the polar gnomonic (central perspective) projection (Snyder, *Map Projections: A Working Manual*, 1987, pp. 164–168), specialized to a spherical Earth and re-centred at p_0; geometrically, it is the perspective projection of S^2_R from its own centre onto the tangent plane at p_0, which is exactly why it is defined only for $\alpha < \pi/2$ (a great circle through p_0 meets the tangent plane at p_0 only within the open hemisphere) and why it sends great circles to straight lines — the property that made it, historically, the natural chart for greatcircle navigation, though not (as this section shows) for controlling distortion. It is the case $\rho_{\text{gn}}(\alpha) := R\tan\alpha$ of Eq. (6.1).

Proposition 7.7 (Gnomonic curvature and the general theorem's hypothesis)**.** *ρ_{gn} is strictly convex on $(0,\pi/2)$, with $\rho'_{\text{gn}}(0) = R$, and satisfies $\rho_{\text{gn}}(x) \leq \rho'_{\text{gn}}(x)\sin x$ for every $x \in (0,\pi/2)$.*

Proof. $\rho'_{\text{gn}}(\alpha) = R\sec^2\alpha$, so $\rho'_{\text{gn}}(0) = R$; $\rho''_{\text{gn}}(\alpha) = 2R\sec^2\alpha\tan\alpha > 0$ for $\alpha \in (0,\pi/2)$, so ρ_{gn} is strictly convex there. The claimed inequality is $\tan x \leq \sec^2 x\sin x = \sin x/\cos^2 x$, equivalent (multiplying by $\cos^2 x > 0$) to $\sin x\cos x \leq \sin x$, i.e., $\cos x \leq 1$, true for every x, with equality only at $x = 0$. □

Theorem 7.8 (Global distortion of the gnomonic projection)**.**

$$D_{\text{gn}}(x) \;=\; \sec^2(x). \tag{7.4}$$

Proof. Immediate from Theorem 7.2(b) and Proposition 7.7: $D_{\text{gn}}(x) = \rho'_{\text{gn}}(x)/R = \sec^2 x$. □

Remark 7.9 (Why the gnomonic projection is so much worse). Table 7.1 below shows $D_{\text{gn}}(x) = \sec^2 x$ is, at every tabulated radius, substantially larger than any of the other three azimuthal projections compared in this monograph, including orthographic projection's own $\sec x$: since $\sec x > 1$ for $x > 0$, $\sec^2 x > \sec x$ always, so the gnomonic projection is provably worse than even the crudest of Chapter 3's benchmarks. This is not a surprise to cartographers — the gnomonic projection's radial profile $R \tan \alpha \to \infty$ as $\alpha \to \pi/2$ is the algebraic signature of its inability to represent a full hemisphere at all, let alone with bounded distortion — but it is worth recording as a caution alongside Section 6.4's more encouraging finding: not every classical, geometrically motivated azimuthal projection is a competitive embedding in the bi-Lipschitz sense of Definition 2.1, and having a clean geometric construction (central perspective, in this case) is no guarantee of good metric behaviour.

7.4 The five-projection comparison, completed

Table 7.1: Global distortion of all five azimuthal projections now considered, together with the universal lower bound. All columns are exact closed forms; the gnomonic column is undefined ($\to \infty$) as $x \to \pi/2$ and is not tabulated beyond the radii already used throughout Part I. Grid-based pipeline verification (`data/general_azimuthal_grid_results.json`) confirms every closed form to a relative error below 2.3×10^{-3} at every radius, including at $x = 1.00$, where Table 6.1's earlier Monte Carlo check had shown larger (1–2%) discrepancies for the equal-area projection specifically; the deterministic grid used here resolves that earlier discrepancy as a finite-sample artefact of random pair sampling, not a flaw in the closed form.

$x = r/R$	$D_{\text{min}}(x)$	$D_{\text{eq}}(x)$	$D_{\text{ea}}(x) = D_{\text{st}}(x)$	$\sec(x)$	$D_{\text{gn}}(x) = \sec^2 x$
0.02	1.000033	1.000067	1.000100	1.000200	1.000400
0.05	1.000208	1.000417	1.000625	1.001251	1.002504
0.10	1.000835	1.001669	1.002504	1.005021	1.010067
0.20	1.003368	1.006698	1.010067	1.020339	1.041091
0.30	1.007677	1.015159	1.022842	1.046752	1.095689
0.50	1.022247	1.042915	1.065199	1.139494	1.298446
0.70	1.046528	1.086589	1.133246	1.307459	1.709450
1.00	1.109574	1.188395	1.298446	1.850816	3.425519

7.4.1 Asymptotic ordering, extended

Expanding $\sec^2 x = 1 + x^2 + O(x^4)$ (from $\sec u = 1 + u^2/2 + O(u^4)$ at $u = x$), the five leading coefficients, in twelfths, are $1, 2, 3, 6, 12$:

$$D_{\text{min}}(x) \;<\; D_{\text{eq}}(x) \;<\; D_{\text{ea}}(x) = D_{\text{st}}(x) \;<\; \sec(x) \;<\; D_{\text{gn}}(x), \qquad x \to 0^+, \tag{7.5}$$

extending Eq. (6.13) and, again, matching Table 7.1's exact values at every tabulated radius rather than only in the small-x limit. The gnomonic projection's coefficient (1, i.e. $12/12$) is exactly twice orthographic projection's ($1/2$, i.e. $6/12$), itself already six times the universal lower bound's ($1/12$); the gnomonic projection is, to leading order, twelve times further above $D_{\text{min}}(x)$ than the azimuthal equidistant projection.

7.5 Local optimality of the equidistant profile

Section 6.5's third open point asked whether some projection — azimuthal or otherwise — could narrow the gap between $D_{\text{eq}}(x)$ and $D_{\min}(x)$ further. Table 7.1 has now shown that three further classical azimuthal projections all do *worse*, not better, than the equidistant one. This section asks a sharper version of the same question: is the equidistant profile merely the best *among the four classical examples tried*, or is it a genuine local minimum of $D_\rho(x)$ within a continuous family of nearby profiles?

Definition 7.10 (Cubic perturbation family)**.** For $p \in \mathbb{R}$, let $\rho_p(\alpha) := R\alpha + pR\alpha^3$, the unique family of profiles obtained by adding, to the equidistant profile $\rho_0(\alpha) = R\alpha$, the lowest-order term consistent with $\rho_p(0) = 0$ and $\rho_p'(0) = R$ (the constant and linear coefficients are fixed by these two conditions; the quadratic coefficient is fixed at 0 since $\rho_p''(0) = 6pR \cdot 0 = 0$ would otherwise make ρ_p change concavity sign at $\alpha = 0$, which Theorem 7.2 does not cover). For $p < 0$, $\rho_p''(\alpha) = 6pR\alpha < 0$ on $(0, x]$: ρ_p is concave; for $p > 0$, convex.

Proposition 7.11 (Asymptotic distortion of the cubic family)**.** *As $x \to 0^+$, for fixed $p \neq 0$,*

$$D_{\rho_p}(x) \;=\; 1 + c(p)\,x^2 + O(x^4), \qquad c(p) \;=\; \begin{cases} \dfrac{1}{6} - 2p, & p \le 0, \\ \max\!\left(3p,\; p + \dfrac{1}{6}\right), & p \ge 0, \end{cases} \tag{7.6}$$

and $c(p) > c(0) = \frac{1}{6}$ for every $p \neq 0$: the coefficient is minimized, uniquely, at $p = 0$.

Proof. For $p < 0$, ρ_p is concave: check the hypothesis of Theorem 7.2(a). $\rho_p(x)/(R\sin x) = (x + px^3)/\sin x$; since $\sin x < x$, and for x small enough that $|p|x^2 < 1$ this ratio exceeds 1 (direct check: $(x+px^3) - \sin x = (x - \sin x) + px^3 = x^3/6 + px^3 + O(x^5) = x^3(1/6 + p) + O(x^5)$, positive for small x whenever $p > -1/6$, which covers all p of interest here), so the hypothesis holds for x small enough. Equation (7.1) then gives $D_{\rho_p}(x) = \rho_p(x)/(\rho_p'(x)\sin x)$. Expand each factor: $\rho_p(x) = R(x + px^3)$; $\rho_p'(x) = R(1 + 3px^2)$; $\sin x = x(1 - x^2/6 + O(x^4))$. Hence

$$D_{\rho_p}(x) = \frac{x + px^3}{(1 + 3px^2)\,x(1 - x^2/6 + O(x^4))} = \frac{1 + px^2}{(1 + 3px^2)(1 - x^2/6 + O(x^4))} = (1+px^2)(1-3px^2)(1+x^2/6)+O(x$$

using $(1+3px^2)^{-1} = 1 - 3px^2 + O(x^4)$; expanding the product to $O(x^2)$, $D_{\rho_p}(x) = 1 + (p - 3p + 1/6)x^2 + O(x^4) = 1 + (1/6 - 2p)x^2 + O(x^4)$, giving $c(p) = 1/6 - 2p$ for $p < 0$, and by continuity at $p = 0$ as well.

For $p > 0$, ρ_p is convex; Proposition 7.7-style direct check confirms $\rho_p(x) \le \rho_p'(x)\sin x$ for x small enough (both sides agree to leading orders x and only the x^3 term can differ, and the required inequality reduces, after the same cancellation as in Proposition 7.7's proof, to a condition satisfied for small x regardless of the sign or size of p), so Theorem 7.2(b) applies whenever it is the binding case, giving $D_{\rho_p}(x) = \rho_p'(x)/R = 1 + 3px^2$ directly, *unless* the tangential supremum $\rho_p(x)/(R\sin x)$ exceeds the radial supremum $\rho_p'(x)/R$ — possible for small $p > 0$, since both are within $O(x^2)$ of 1. Comparing the two candidate suprema's expansions, $\rho_p(x)/(R\sin x) = 1 + (p + 1/6)x^2 + O(x^4)$ (identical computation to the concave case above, since this expansion did not use the sign of p) against $\rho_p'(x)/R = 1 + 3px^2$: the true supremum, and hence $D_{\rho_p}(x)$'s leading coefficient, is $\max(3p,\, p + 1/6)$, matching Eq. (7.6).

Finally, $c(p) > 1/6$ for every $p \neq 0$: for $p < 0$, $c(p) = 1/6 - 2p > 1/6$ since $-2p > 0$; for $p > 0$, $c(p) = \max(3p, p + 1/6) \ge p + 1/6 > 1/6$. Equality $c(p) = 1/6$ holds only at $p = 0$. □

Table 7.2 corroborates Eq. (7.6) against the same deterministic grid pipeline used in Table 7.1, at two fixed radii, across a symmetric range of perturbation strengths p.

Table 7.2: Grid-computed $D_{\rho_p}(x)$ for the cubic perturbation family (Definition 7.10) at $x = 0.3$ and $x = 0.5$, alongside the predicted asymptotic excess coefficient $(D_{\rho_p}(x) - 1)/x^2$ and Eq. (7.6)'s closed-form prediction $c(p)$. The minimum in each column of excess coefficients occurs at $p = 0$, as predicted; the growing gap between the two rightmost columns at $x = 0.5$ versus $x = 0.3$ reflects expected $O(x^2)$ higher-order corrections to the leading asymptotic, not a failure of Proposition 7.11. Source: `code/cubic_perturbation_check.py`, `data/cubic_perturbation_results.json`.

p	$D_{\rho_p}(0.3)$	$(D-1)/x^2$	$c(p)$ pred.	$D_{\rho_p}(0.5)$	$(D-1)/x^2$
−0.10	1.033891	0.376567	0.366667	1.099140	0.396560
−0.05	1.024397	0.271079	0.266667	1.069934	0.279736
−0.02	1.018824	0.209158	0.206667	1.053476	0.213904
−0.01	1.016987	0.188741	0.186667	1.048156	0.192622
0.00	1.015159	0.168433	0.166667	1.042915	0.171659
0.01	1.016073	0.178585	0.176667	1.045522	0.182088
0.02	1.016986	0.188737	0.186667	1.048129	0.192518
0.05	1.019727	0.219191	0.216667	1.055951	0.223805
0.10	1.026955	0.299499	0.300000	1.074875	0.299499

Key Result

Within the cubic perturbation family of Definition 7.10 — the natural one-parameter family of smooth deformations of the equidistant profile consistent with the centre-normalization $\rho(0) = 0, \rho'(0) = R$ — the azimuthal equidistant projection ($p = 0$) is a strict local minimum of global distortion, to leading asymptotic order, confirmed both by closed-form Taylor expansion (Proposition 7.11) and by independent deterministic grid computation (Table 7.2).

7.6 What this chapter does, and does not, establish

Theorem 7.2 unifies Chapter 6's two case-by-case derivations into a single argument depending only on the sign of ρ'', subject to an explicit auxiliary hypothesis verified separately for each profile to which it is applied (Corollaries 7.3 and 7.4 and Proposition 7.7); Section 7.3 used it to add a fifth classical azimuthal projection, confirming it is markedly worse than the other four; and Section 7.5 used it to prove that the azimuthal equidistant projection is a strict local minimum of distortion within one natural, low-dimensional family of perturbations. Several points remain explicitly open:

1. **The auxiliary hypotheses of Theorem 7.2 are not shown to hold for every concave or convex profile.** Only the three profiles of Sections 7.2 and 7.3 (and the cubic family of Section 7.5, checked directly in Proposition 7.11's proof) are verified; a concave or convex profile violating either hypothesis would require a separate, case-specific argument, not covered here.

2. **Local optimality is proved only within the cubic family, and only to leading asymptotic order.** Proposition 7.11 does not address quartic or higher-order perturbations, nor a fully general Gateaux-type first-variation argument over all smooth perturbations $\rho = R\alpha + \varepsilon\eta(\alpha)$ with $\eta(0) = \eta'(0) = 0$ for arbitrary η; such a fully general local-optimality theorem would need to handle the non-smoothness of the sup / inf defining $D_\rho(x)$ directly, via an envelope-theorem argument rather than the direct Taylor expansion used here,

which exploited the cubic family's simplicity. Chapter 8 carries out exactly this argument: it shows the required first-order sign pattern holds for *every* C^1 perturbation direction, recovering this chapter's cubic-family result as a special case, though it leaves strictness in certain degenerate directions, and second-order effects generally, still open (Section 8.6).

3. **Local optimality is not global optimality.** Even a fully general local result would leave open whether some structurally different, non-perturbative embedding — azimuthal or not — achieves distortion strictly between $D_{\text{eq}}(x)$ and $D_{\min}(x)$; Section 5.6's original question, whether $D_{\text{eq}}(x) = D^{\star}(x)$ exactly, is narrowed by this chapter's local result but not answered.

Chapter 8

A General First-Variation Theorem: The Equidistant Profile as a Gateaux-Stationary Point

Why this chapter was added

Section 7.6's second open point was explicit: Proposition 7.11 proves local optimality of the azimuthal equidistant profile only within the one-parameter cubic family $\rho_p(\alpha) = R\alpha + pR\alpha^3$, and only via a direct Taylor expansion that exploits that family's simplicity; a fully general first-variation argument, covering an arbitrary smooth perturbation $\rho_\varepsilon(\alpha) = R\alpha + \varepsilon\eta(\alpha)$ with $\eta(0) = \eta'(0) = 0$, was left open, with the added remark that such an argument would have to confront the non-smoothness of the sup / inf defining $D_\rho(x)$ directly, via an envelope-theorem (Danskin-type) argument rather than a direct expansion. This chapter carries out exactly that argument. The conclusion is a genuine generalization: for *every* admissible perturbation direction η (not merely cubic ones), the one-sided Gateaux derivatives of $D_{\rho_\varepsilon}(x)$ at $\varepsilon = 0$ have the sign required for $\varepsilon = 0$ to be a first-order local minimum (Theorem 8.8); the cubic family's sharper, strict, second-order conclusion (Proposition 7.11) is recovered as a special case (Remark 8.9), and the general theorem is checked, independently, against the same deterministic grid pipeline used throughout this monograph, on perturbations genuinely outside the cubic family (Section 8.5).

8.1 Setup: an arbitrary smooth perturbation of the equidistant profile

Fix a cap half-angle $x \in (0, \pi/2)$ throughout this chapter. Let $\eta \in C^1([0, x])$ satisfy

$$\eta(0) = 0, \qquad \eta'(0) = 0, \tag{8.1}$$

the same centre-normalization constraint already imposed on every profile in Chapter 7 (so that $\rho_\varepsilon(0) = 0$ and $\rho'_\varepsilon(0) = R$ for every ε), and set

$$\rho_\varepsilon(\alpha) \;:=\; R\alpha + \varepsilon\,\eta(\alpha), \qquad \alpha \in [0, x]. \tag{8.2}$$

No sign or convexity condition is imposed on η: unlike Chapter 7's theorem, which required ρ to have one fixed sign of curvature throughout, η here is an *arbitrary* C^1 function subject only to Eq. (8.1). In particular ρ''_ε may change sign on $(0, x]$, so Theorem 7.2 does not apply to ρ_ε directly for $\varepsilon \neq 0$, and $D_{\rho_\varepsilon}(x)$ has no closed form available in general; this chapter studies its behaviour only in a neighbourhood of $\varepsilon = 0$, where $\rho_0(\alpha) = R\alpha$ is the azimuthal equidistant profile itself.

Throughout, write, for a radial pair $0 < \alpha_1 < \alpha_2 \le x$ (common azimuth) and a tangential pair $\alpha \in (0, x]$, $t = \sin(\Delta\theta/2) \in (0, 1]$ (Chapter 5's parametrization, via the identity $d_{S^2_R}(q_1, q_2) =$

$2R\arcsin(t\sin\alpha)$ already used there and reused unchanged in Chapter 7), the two families of distance ratios

$$f_{\text{rad}}(\alpha_1,\alpha_2;\varepsilon) := \frac{\rho_\varepsilon(\alpha_2)-\rho_\varepsilon(\alpha_1)}{R(\alpha_2-\alpha_1)}, \qquad f_{\tan}(\alpha,t;\varepsilon) := \frac{\rho_\varepsilon(\alpha)\,t}{R\arcsin(t\sin\alpha)}, \tag{8.3}$$

so that $D_{\rho_\varepsilon}(x) = S(\varepsilon)/I(\varepsilon)$ with $S(\varepsilon) := \max\big(S_{\text{rad}}(\varepsilon), S_{\tan}(\varepsilon)\big)$, $I(\varepsilon) := \min\big(I_{\text{rad}}(\varepsilon), I_{\tan}(\varepsilon)\big)$, where $S_{\text{rad}}, I_{\text{rad}}$ are the supremum and infimum of f_{rad} over all radial pairs and $S_{\tan}, I_{\tan}$ those of $f_{\tan}$ over all tangential pairs, exactly as in Chapters 5 and 7 but now written with ε made explicit. At $\varepsilon = 0$, $f_{\text{rad}}(\cdot;0) \equiv 1$ (Proposition 5.2) and $f_{\tan}(\alpha,t;0) = \alpha t/\arcsin(t\sin\alpha)$ ranges over $[1, x/\sin x]$ (Theorem 5.6), so $S(0) = x/\sin x$, $I(0) = 1$, recovering Eq. (5.6).

A standing regularity assumption

The argument below differentiates $S(\varepsilon)$ and $I(\varepsilon)$ at $\varepsilon = 0$ by differentiating under a supremum or infimum taken over a fixed (compact, or compactified by the smooth boundary extension of Lemma 8.2 below) parameter domain. This is legitimate whenever the family $\{f(\cdot;\varepsilon)\}$ is, in ε, uniformly (in the parameter) C^1 near $\varepsilon = 0$ with a jointly continuous ε-derivative — which holds here because $\eta \in C^1([0,x])$ and $[0,x]$, $[0,1]$ are compact, so η and η' are bounded and the elementary quotients in Eq. (8.3) depend on ε affinely with parameter-uniformly bounded coefficients. No further regularity is assumed of η beyond $C^1([0,x])$.

8.2 The supremum branch: a non-degenerate corner maximizer

Lemma 8.1 (Strict separation of the two supremum branches). *$S_{\text{rad}}(0) = 1 < x/\sin x = S_{\tan}(0)$, and both $S_{\text{rad}}(\varepsilon)$, $S_{\tan}(\varepsilon)$ are continuous at $\varepsilon = 0$. Consequently there is $\varepsilon_0 > 0$ such that $S(\varepsilon) = S_{\tan}(\varepsilon)$ for all $|\varepsilon| < \varepsilon_0$.*

Proof. $S_{\text{rad}}(0) = 1$ since $f_{\text{rad}}(\cdot;0) \equiv 1$ (Proposition 5.2); $S_{\tan}(0) = x/\sin x > 1$ for $x \in (0,\pi/2)$ by Theorem 5.6 together with $x/\sin x > 1$ there. Continuity of $S_{\text{rad}}, S_{\tan}$ at $\varepsilon = 0$ follows from the standing regularity assumption: a supremum over a parameter set of a family jointly continuous in (parameter, ε), with ε-dependence affine and parameter-uniformly bounded, is itself continuous in ε (a standard fact — $|S_\bullet(\varepsilon) - S_\bullet(0)| \le \sup_{\text{pairs}} |f(\cdot;\varepsilon) - f(\cdot;0)| \to 0$ as $\varepsilon \to 0$, using the uniform bound on η, η'). Given strict separation $S_{\text{rad}}(0) < S_{\tan}(0)$ and continuity of both, there is $\varepsilon_0 > 0$ with $S_{\text{rad}}(\varepsilon) < S_{\tan}(\varepsilon)$ for $|\varepsilon| < \varepsilon_0$, so the max is $S_{\tan}(\varepsilon)$ throughout that range. □

Lemma 8.2 (The tangential supremum is a smooth corner maximum). *$S_{\tan}(\varepsilon)$ is differentiable (two-sided) at $\varepsilon = 0$, with*

$$S'_{\tan}(0) \;=\; \frac{\eta(x)}{R\sin x}. \tag{8.4}$$

Proof. Extend $g(\alpha,t) := f_{\tan}(\alpha,t;0) = \alpha t/\arcsin(t\sin\alpha)$, originally defined for $t \in (0,1]$, to $t = 0$ by $g(\alpha,0) := \alpha/\sin\alpha$ (the value of Lemma 5.4's limit). Expanding $\arcsin(t\sin\alpha) = t\sin\alpha + (t\sin\alpha)^3/6 + O(t^5)$ as $t \to 0$ gives

$$g(\alpha,t) = \frac{\alpha}{\sin\alpha}\Big(1 - \tfrac{1}{6}(t\sin\alpha)^2 + O(t^4)\Big) = \frac{\alpha}{\sin\alpha} - \frac{\alpha\sin\alpha}{6}\,t^2 + O(t^4),$$

so g extends to a function of $(\alpha,t) \in (0,x]\times[0,1]$ that is even and C^1 (indeed smooth) in t at $t = 0$, with $\partial g/\partial t(\alpha,0) = 0$. The same expansion applied to $f_{\tan}(\alpha,t;\varepsilon)$ for small ε shows

the ε-derivative extends continuously to $t = 0$ as well, with $\partial_\varepsilon f_{\tan}(\alpha, 0; 0) = \eta(\alpha)/(R \sin\alpha)$ (set $t = 0$ directly in $\partial_\varepsilon f_{\tan} = \eta(\alpha)t/(R \arcsin(t \sin\alpha))$ after cancelling the common factor $t/\arcsin(t\sin\alpha) \to 1/\sin\alpha$).

On the resulting domain $(\alpha, t) \in (0, x] \times [0, 1]$, $g(\alpha, 0) = \alpha/\sin\alpha$ is strictly increasing in α (its derivative is $(\sin\alpha - \alpha\cos\alpha)/\sin^2\alpha > 0$ for $\alpha \in (0, \pi)$, since $\tan\alpha > \alpha$ there), so its maximum over $\alpha \in (0, x]$ is attained uniquely at $\alpha = x$; and *Theorem* 5.6's own upper bound, $g(\alpha, t) \le \alpha/\sin\alpha \le x/\sin x$ for all (α, t), with equality only at $(\alpha, t) = (x, 0)$, shows $(x, 0)$ is the *unique* maximizer of g over the whole compact domain $(0, x] \times [0, 1]$ (closure at $\alpha \to 0$ poses no difficulty, since $g(\alpha, 0) \to 1 < x/\sin x$ there). A unique maximizer of a family jointly C^1 in (α, t, ε) over a domain compactified exactly at that maximizer is the classical (non-degenerate) case of the envelope theorem: since there is no competing near-maximizer to create one-sided asymmetry, $S_{\tan}(\varepsilon) = f_{\tan}(x, 0; \varepsilon) + o(\varepsilon)$, i.e. $S_{\tan}$ is differentiable at $\varepsilon = 0$ (both one-sided derivatives agree) with derivative equal to $\partial_\varepsilon f_{\tan}(x, 0; 0) = \eta(x)/(R \sin x)$, which is Eq. (8.4). □

Corollary 8.3 (Derivative of the overall supremum). *$S(\varepsilon)$ is differentiable at $\varepsilon = 0$ with $S'(0) = \eta(x)/(R\sin x)$, and $S(0) = x/\sin x$.*

Proof. Immediate from Lemmas 8.1 and 8.2: $S \equiv S_{\tan}$ on a neighbourhood of $\varepsilon = 0$. □

8.3 The infimum branch: a genuinely degenerate tie

Unlike the supremum, the infimum at $\varepsilon = 0$ is *not* a strict, isolated minimum of one branch over the other: both the radial and tangential branches attain the value 1 on entire continua.

Lemma 8.4 (Tied, degenerate infima). *$I_{\text{rad}}(0) = 1$, attained by* every *radial pair; $I_{\tan}(0) = 1$, attained by* every *tangential pair with $t = 1$ (i.e. $\Delta\theta = \pi$).*

Proof. $I_{\text{rad}}(0) = 1$ is Proposition 5.2 again ($f_{\text{rad}}(\cdot; 0) \equiv 1$ identically). For $I_{\tan}(0)$: at $t = 1$, $g(\alpha, 1) = \alpha/\arcsin(\sin\alpha) = \alpha/\alpha = 1$ for every $\alpha \in (0, x] \subset (0, \pi/2)$ (using $\arcsin(\sin\alpha) = \alpha$ on this range), and Theorem 5.6's lower bound gives $g(\alpha, t) \ge 1$ everywhere, so $t = 1$ is exactly where this lower bound is attained, for every α. □

Each branch's one-sided derivative is itself given by a Danskin formula over its own (degenerate) active set.

Lemma 8.5 (One-sided derivatives of the radial infimum branch). *Let $m_1 := \min_{\alpha\in[0,x]} \eta'(\alpha)/R$ and $M_1 := \max_{\alpha\in[0,x]} \eta'(\alpha)/R$ (well-defined since η' is continuous on the compact interval $[0, x]$). Then I_{rad} has right- and left-hand derivatives at $\varepsilon = 0$ equal to*

$$\big(I_{\text{rad}}\big)'_+(0) = m_1, \qquad \big(I_{\text{rad}}\big)'_-(0) = M_1.$$

Proof. By the mean value theorem, for any $0 < \alpha_1 < \alpha_2 \le x$, $\partial_\varepsilon f_{\text{rad}}(\alpha_1, \alpha_2; 0) = \big(\eta(\alpha_2) - \eta(\alpha_1)\big)/\big(R(\alpha_2 - \alpha_1)\big) = \eta'(c)/R$ for some $c \in (\alpha_1, \alpha_2)$; conversely, for any $c \in (0, x)$, letting $\alpha_1, \alpha_2 \to c$ realizes $\eta'(c)/R$ as a limit of such secant slopes, and letting $\alpha_1 \to 0, \alpha_2 \to 0$ realizes $\eta'(0)/R = 0$ in the limit. Hence the closure of $\{\partial_\varepsilon f_{\text{rad}}(\alpha_1, \alpha_2; 0) : 0 < \alpha_1 < \alpha_2 \le x\}$ equals $[m_1, M_1]$ exactly (the continuous image, under the secant-slope map, of the connected compact set of pairs, using continuity of η' on $[0, x]$). Since *every* pair (α_1, α_2) is in the active set for $I_{\text{rad}}(0) = 1$ (Lemma 8.4), the Danskin envelope theorem for an infimum over an active set $\Theta^\star$ of a family affine in ε gives right-derivative $\inf_{\Theta^\star} \partial_\varepsilon f = m_1$ and left-derivative $\sup_{\Theta^\star} \partial_\varepsilon f =$

M_1 (the two-sided generalization of the standard one-sided Danskin formula, obtained by applying the one-sided statement to $\varepsilon \mapsto \varepsilon$ and to $\varepsilon \mapsto -\varepsilon$ in turn). □

Lemma 8.6 (One-sided derivatives of the tangential infimum branch). *Let $\widetilde{\eta}(\alpha) := \eta(\alpha)/\alpha$ for $\alpha \in (0, x]$, extended to $\widetilde{\eta}(0) := \eta'(0) = 0$ (continuous by $\eta \in C^1$, L'Hôpital). Let $m_2 := \min_{\alpha\in[0,x]} \widetilde{\eta}(\alpha)/R$, $M_2 := \max_{\alpha\in[0,x]} \widetilde{\eta}(\alpha)/R$. Then*

$$\big(I_{\tan}\big)'_+(0) = m_2, \qquad \big(I_{\tan}\big)'_-(0) = M_2.$$

Proof. By Lemma 8.4, the active set for $I_{\tan}(0) = 1$ is exactly $\{(\alpha, 1) : \alpha \in (0, x]\}$, on which $\partial_\varepsilon f_{\tan}(\alpha, 1; 0) = \eta(\alpha)\cdot 1/(R \arcsin(\sin\alpha)) = \eta(\alpha)/(R\alpha) = \widetilde{\eta}(\alpha)/R$. The Danskin envelope theorem for the infimum over this active set (a continuous image of the compact interval $(0, x]$, compactified at $\alpha \to 0$ by the extension $\widetilde{\eta}(0) = 0$, exactly as in Lemma 8.5) gives right-derivative $\inf \widetilde{\eta}/R = m_2$ and left-derivative $\sup \widetilde{\eta}/R = M_2$. □

Proposition 8.7 (Derivative of the overall infimum). *$I(\varepsilon) = \min(I_{\text{rad}}(\varepsilon), I_{\tan}(\varepsilon))$ has one-sided derivatives at $\varepsilon = 0$*

$$I'_+(0) = \min(m_1, m_2), \qquad I'_-(0) = \max(M_1, M_2).$$

Proof. Both $I_{\text{rad}}, I_{\tan}$ equal 1 at $\varepsilon = 0$ (Lemma 8.4) and each has one-sided derivatives there (Lemmas 8.5 and 8.6). For finitely many functions tied at a point and each one-sided differentiable there, the pointwise minimum's one-sided derivatives are the minima (respectively) of the individual one-sided derivatives on the corresponding side — a direct consequence of the definition of one-sided derivative applied to $\min(a(\varepsilon), b(\varepsilon)) = a(0) + \min\big(a(\varepsilon) - a(0),\ b(\varepsilon) - b(0)\big)$ and passing to the limit termwise, no further active-set subtlety being needed since there are only two branches. This gives $I'_+(0) = \min\big((I_{\text{rad}})'_+(0), (I_{\tan})'_+(0)\big) = \min(m_1, m_2)$ and likewise $I'_-(0) = \max(M_1, M_2)$. □

8.4 The general local-optimality theorem

Theorem 8.8 (Gateaux-stationarity of the azimuthal equidistant profile). *For every $x \in (0, \pi/2)$ and every $\eta \in C^1([0, x])$ with $\eta(0) = \eta'(0) = 0$, the function $\varepsilon \mapsto D_{\rho_\varepsilon}(x)$ has one-sided derivatives at $\varepsilon = 0$ satisfying*

$$D'_+(0) \ \geq\ 0, \qquad D'_-(0) \ \leq\ 0. \tag{8.5}$$

Consequently $\varepsilon = 0$ — the azimuthal equidistant profile itself — is a first-order local minimum of $D_{\rho_\varepsilon}(x)$ along every admissible perturbation direction η: to first order in ε, no smooth perturbation of the equidistant profile, in either direction, can decrease its global distortion.

Proof. Since S is differentiable at 0 (Corollary 8.3) and $I(0) = 1 \neq 0$, the quotient $D = S/I$ has one-sided derivatives at 0 given by the quotient rule applied on each side separately,

$$D'_+(0) = S'(0)\cdot I(0) - S(0)\cdot I'_+(0) = \frac{\eta(x)}{R\sin x} - \frac{x}{\sin x}\min(m_1, m_2),$$

$$D'_-(0) = S'(0)\cdot I(0) - S(0)\cdot I'_-(0) = \frac{\eta(x)}{R\sin x} - \frac{x}{\sin x}\max(M_1, M_2),$$

using $I(0) = 1$, Corollary 8.3's value of $S'(0)$, Proposition 8.7's values of $I'_\pm(0)$, and $S(0) = x/\sin x$. Multiplying through by $R\sin x > 0$, the two claimed inequalities are equivalent to

$$\eta(x) \ \geq\ Rx\cdot\min(m_1, m_2), \qquad \eta(x) \ \leq\ Rx\cdot\max(M_1, M_2). \tag{8.6}$$

Both are immediate from the definitions of m_2, M_2 as, respectively, the minimum and maximum of $\widetilde{\eta}(\alpha)/R = \eta(\alpha)/(R\alpha)$ over $\alpha \in [0, x]$ (Lemma 8.6), which in particular bound the value $\widetilde{\eta}(x)/R = \eta(x)/(Rx)$ attained at the right endpoint $\alpha = x \in [0, x]$:

$$m_2 \;\le\; \frac{\eta(x)}{Rx} \;\le\; M_2.$$

Since $\min(m_1, m_2) \le m_2$ and $\max(M_1, M_2) \ge M_2$, multiplying the two displayed inequalities by $Rx > 0$ gives exactly Eq. (8.6), i.e. Eq. (8.5). □

What this theorem does and does not claim

Theorem 8.8 is deliberately stated with non-strict inequalities. The proof shows $D'_+(0) \ge 0$ using only $\min(m_1, m_2) \le m_2 \le \eta(x)/(Rx)$, with equality possible precisely when the minimum of $\widetilde{\eta}$ over $[0, x]$ is attained at the right endpoint $\alpha = x$ itself (and $m_1 \ge m_2$, so the radial branch does not push the minimum lower); in that case $D'_+(0) = 0$ and the first-order argument is silent on whether $\varepsilon = 0$ is a strict local minimum in that particular direction — exactly the situation Section 7.6's point 2 anticipated in noting that a general argument would need to go beyond the direct expansion used for the cubic family specifically. Resolving strictness in every direction, or in the degenerate directions where Eq. (8.5) holds with equality, would require a second-order (Hessian-level) version of this same envelope argument, which is not undertaken here.

Remark 8.9 (Recovering the cubic family). Taking $\eta(\alpha) = \alpha^3$ (so that $\rho_\varepsilon = \rho_p$ of Definition 7.10 with $\varepsilon = p$, $R = 1$) gives $\eta'(\alpha) = 3\alpha^2$, hence $m_1 = 0 = M_1$'s lower endpoint value at $\alpha = 0$ but $M_1 = 3x^2$ (attained at $\alpha = x$); $\widetilde{\eta}(\alpha) = \alpha^2$, so $m_2 = 0$ (at $\alpha = 0$), $M_2 = x^2$ (at $\alpha = x$). Hence $\min(m_1, m_2) = 0$ and $\max(M_1, M_2) = 3x^2$, giving

$$D'_+(0) = \frac{x^3}{\sin x} - \frac{x}{\sin x}\cdot 0 = \frac{x^3}{\sin x}, \qquad D'_-(0) = \frac{x^3}{\sin x} - \frac{x}{\sin x}\cdot 3x^2 = -\frac{2x^3}{\sin x}.$$

These have exactly the ratio $2:1$ between the two one-sided slopes already visible in Proposition 7.11's asymptotic coefficient $c(p)$ (whose one-sided slopes in p at $p = 0$ are -2 for $p < 0$ and $+1$ for $p > 0$, the same $2:1$ ratio, there expressed as a coefficient of the small-x expansion rather than directly as a p-derivative at fixed x): the two computations are answering different but consistent questions, and agree wherever they overlap. This chapter's theorem is strictly more general, covering every C^1 direction η at every fixed x, not only the cubic family's leading small-x asymptotics.

8.5 Numerical verification against the grid pipeline

Theorem 8.8's closed-form one-sided derivatives, and the sign pattern Eq. (8.5), are checked here against the same deterministic grid distortion pipeline used throughout Chapter 7 (`code/gateaux_general_perturbation.py`), on four perturbation directions chosen specifically to lie *outside* the cubic family: a positive and a negative pure quartic ($\eta(\alpha) = \pm\alpha^4$), a transcendental direction ($\eta(\alpha) = \sin\alpha - \alpha$, which is negative and non-polynomial), and a mixed cubic-quartic direction ($\eta(\alpha) = \alpha^3 - \alpha^4$, chosen so that $\widetilde{\eta}$ is non-monotonic on $[0, x]$). For each direction and each of two cap radii, the script computes $D_{\rho_\varepsilon}(x)$ by direct grid search over radial and tangential pairs (identical methodology to `code/cubic_perturbation_check.py`) at $\varepsilon = \pm 0.02, \pm 0.01, \pm 0.005$, forms the finite-difference quotients $\big(D(\varepsilon) - D(0)\big)/\varepsilon$ as one-sided approximations to $D'_\pm(0)$, and compares them against the closed-form predictions of Eq. (8.6)'s underlying quotient-rule formulas.

Table 8.1: Closed-form one-sided Gateaux derivative predictions (Theorem 8.8's proof) against finite-difference estimates from the grid pipeline at $\varepsilon = \pm 0.005$ (the smallest step tested, hence the closest finite-difference approximation to the true one-sided derivative). All eight rows independently confirm the predicted sign pattern $D'_+(0) \geq 0$, $D'_-(0) \leq 0$. Source: `code/gateaux_general_perturbation.py`, `data/gateaux_general_perturbation_results.json`.

$\eta(\alpha)$	x	$D'_+(0)$ pred.	FD, ε=0.005	$D'_-(0)$ pred.	FD, ε=−0.005
α^4	0.3	0.027409	0.027409	−0.082228	−0.082155
α^4	0.5	0.130364	0.130364	−0.391093	−0.391512
$-\alpha^4$	0.3	0.082228	0.082155	−0.027409	−0.027409
$-\alpha^4$	0.5	0.391093	0.391512	−0.130364	−0.130364
$\sin\alpha - \alpha$	0.3	0.030182	0.030156	−0.015159	−0.015159
$\sin\alpha - \alpha$	0.5	0.084756	0.084719	−0.042915	−0.042915
$\alpha^3 - \alpha^4$	0.3	0.063955	0.063955	−0.100501	−0.100504
$\alpha^3 - \alpha^4$	0.5	0.130364	0.130364	−0.130364	−0.130527

Key Result

Across all four tested perturbation directions — two of them (the quartics) sign-definite but of a different order than the cubic family, one transcendental, and one deliberately non-monotonic in $\widetilde{\eta}$ — and both tested cap radii, the closed-form predictions of Theorem 8.8's proof agree with independent finite-difference grid computation to within 10^{-3} or better (consistent with the expected $O(\varepsilon)$ truncation error of a one-sided finite difference at $\varepsilon = 0.005$), and every predicted sign matches $D'_+(0) \geq 0$, $D'_-(0) \leq 0$ exactly as Eq. (8.5) requires. None of these four directions lies in the cubic family of Definition 7.10, so this is a genuinely independent check of the general theorem, not a re-verification of Proposition 7.11.

8.6 What this chapter does, and does not, establish

This chapter closes the specific gap identified in Section 7.6, point 2: local optimality of the azimuthal equidistant profile, at first order, is now established for *every* smooth perturbation direction, not only the cubic family, via a Danskin/envelope-theorem argument that explicitly confronts (rather than avoids) the non-smoothness of the sup / inf defining $D_\rho(x)$ — carried out separately for the supremum branch (Section 8.2, a non-degenerate corner maximizer, differentiable in the ordinary two-sided sense) and the infimum branch (Section 8.3, a genuinely degenerate tie between two continua, requiring the full one-sided Danskin formula). Three points remain open:

1. **Strictness in degenerate directions.** As noted after Theorem 8.8, Eq. (8.5) can hold with equality for specific η (those whose $\widetilde{\eta}$ attains its extremum exactly at $\alpha = x$); a second-order envelope argument, differentiating $I'_\pm(0)$ itself, would be needed to resolve strict local optimality in those directions and is not carried out here. *Closed in Chapter 9: no such η exists among nontrivial perturbations, so no second-order argument is needed after all.*

2. **Non-C^1 perturbations.** The standing regularity assumption requires $\eta \in C^1([0, x])$; whether the same conclusion extends to merely Lipschitz or one-sided-differentiable η (relevant, e.g., to perturbations with a kink) is not addressed. *Closed in Chapter 9, Section 9.2: the argument survives the passage to piecewise-C^1 (kinked) η essentially unchanged.*

3. **Local optimality is still not global optimality.** Exactly as Section 7.6's point 3 already noted for the cubic family specifically, this chapter's much more general first-order result still says nothing about whether some non-perturbative, structurally different embedding beats $D_{\text{eq}}(x)$; Section 5.6's original question, $D_{\text{eq}}(x) \stackrel{?}{=} D^{\star}(x)$, remains exactly as open as before.

Chapter 9

Strict Gateaux Optimality: Closing the Degenerate-Direction Gap, and Kinked Perturbations

Why this chapter was added

Section 8.6 left two specific points open after Theorem 8.8: first, that the theorem's inequalities Eq. (8.5) are stated non-strictly, since the proof only shows $D'_+(0) \geq 0$ via $\min(m_1, m_2) \leq m_2 \leq \eta(x)/(Rx)$, which can hold with equality; second, that the standing regularity assumption $\eta \in C^1([0, x])$ leaves merely Lipschitz or kinked perturbations unaddressed. This chapter closes both points completely. Section 9.1 shows that the degenerate case — equality in either half of Eq. (8.5) — forces $\eta \equiv 0$ on $[0, x]$: among *nontrivial* perturbations, Eq. (8.5) always holds strictly, so no second-order argument is needed to resolve strictness, because there is nothing left to resolve. Section 9.2 then shows the entire argument of Chapter 8, including this chapter's strictness conclusion, survives unchanged when η is only piecewise-C^1 (finitely many kinks), provided the one-sided derivatives at each kink are included alongside the smooth-piece values when forming m_1, M_1. Both conclusions are checked numerically in Section 9.3 against the same deterministic grid pipeline used throughout this monograph.

9.1 Degenerate directions do not exist among nontrivial perturbations

Recall from Theorem 8.8's proof that, writing $k := \eta(x)/(Rx) = \widetilde{\eta}(x)/R$,

$$D'_+(0) = \frac{x}{\sin x}\Big(k - \min(m_1, m_2)\Big), \qquad D'_-(0) = \frac{x}{\sin x}\Big(k - \max(M_1, M_2)\Big), \tag{9.1}$$

where m_1, M_1 are the min and max of η'/R over $[0, x]$ (Lemma 8.5) and m_2, M_2 are the min and max of $\widetilde{\eta}/R = \eta(\alpha)/(R\alpha)$ over $[0, x]$ (Lemma 8.6, with $\widetilde{\eta}(0) := \eta'(0) = 0$). Since $m_2 \leq k \leq M_2$ trivially (as $k = \widetilde{\eta}(x)/R$ is itself one of the values over which m_2, M_2 range), $D'_+(0) \geq 0$ and $D'_-(0) \leq 0$ always, recovering Eq. (8.5); the question is exactly when equality holds.

Lemma 9.1 (Equality in the right-derivative forces $m_1 \geq m_2 = k$)**.** *$D'_+(0) = 0$ if and only if both*

$$(A)\ \ \widetilde{\eta}(\alpha) \geq \widetilde{\eta}(x)\ \ \forall\, \alpha \in [0, x], \qquad (B)\ \ \eta'(\alpha) \geq \eta(x)/x\ \ \forall\, \alpha \in [0, x]. \tag{9.2}$$

Proof. By Eq. (9.1), $D'_+(0) = 0 \iff \min(m_1, m_2) = k$. Since $m_2 = \min_\alpha \widetilde{\eta}(\alpha)/R \leq \widetilde{\eta}(x)/R = k$ always (the minimum over $[0, x]$ cannot exceed the value at the particular point $\alpha = x$), $\min(m_1, m_2) \leq m_2 \leq k$, so $\min(m_1, m_2) = k$ forces both $m_2 = k$ — which is exactly (A), $\widetilde{\eta}$ attaining its minimum over $[0, x]$ at $\alpha = x$ itself — and $m_1 \geq m_2$ (else $\min(m_1, m_2) = m_1 <$

$m_2 = k$, contradicting $\min(m_1, m_2) = k$). The condition $m_1 \geq m_2 = k$, i.e. $\min_\alpha \eta'(\alpha)/R \geq k$, is exactly (B). Conversely, (A) and (B) together give $m_2 = k$ and $m_1 \geq k = m_2$, hence $\min(m_1, m_2) = m_2 = k$, i.e. $D'_+(0) = 0$. □

Theorem 9.2 (No nontrivial degenerate direction)**.** *If $\eta \in C^1([0, x])$ satisfies $\eta(0) = \eta'(0) = 0$ and $D'_+(0) = 0$, then $\eta \equiv 0$ on $[0, x]$. Symmetrically, if $D'_-(0) = 0$, then $\eta \equiv 0$ on $[0, x]$. Consequently, for every $\eta \not\equiv 0$,*

$$D'_+(0) \;>\; 0, \qquad D'_-(0) \;<\; 0, \tag{9.3}$$

and $\varepsilon = 0$ is a strict *first-order local minimum of $D_{\rho_\varepsilon}(x)$ along every nontrivial perturbation direction.*

Proof. Suppose $D'_+(0) = 0$; by Lemma 9.1, condition (B) gives $\eta'(\alpha) \geq k$ for every $\alpha \in [0, x]$, where $k = \eta(x)/x$ is a fixed constant. Integrating this pointwise inequality over $[0, x]$ and using $\eta(0) = 0$,

$$\eta(x) \;=\; \eta(x) - \eta(0) \;=\; \int_0^x \eta'(\alpha)\, d\alpha \;\geq\; \int_0^x k\, d\alpha \;=\; kx \;=\; \eta(x),$$

so equality holds throughout: $\int_0^x \big(\eta'(\alpha) - k\big)\, d\alpha = 0$ with the integrand continuous and ≥ 0 everywhere on $[0, x]$ by (B), which forces $\eta'(\alpha) = k$ for *every* $\alpha \in [0, x]$ (a continuous nonnegative function with zero integral over an interval vanishes identically on that interval). Hence $\eta' \equiv k$, so $\eta(\alpha) = k\alpha$ (using $\eta(0) = 0$). But $\eta'(0) = 0$ is required by Eq. (8.1), and $\eta'(0) = k$ from the just-derived identity, so $k = 0$; thus $\eta \equiv 0$.

The case $D'_-(0) = 0$ is symmetric: Lemma 9.1's argument, with max in place of min and M_1, M_2 in place of m_1, m_2, gives $\eta'(\alpha) \leq k$ for every α, and the same integration argument (with the inequality reversed) forces $\eta' \equiv k$, hence again $k = 0$ and $\eta \equiv 0$.

Since $D'_+(0) \geq 0$ and $D'_-(0) \leq 0$ always (Theorem 8.8), and either equality forces $\eta \equiv 0$, every $\eta \not\equiv 0$ gives $D'_+(0) > 0$ and $D'_-(0) < 0$ strictly, i.e. Eq. (9.3). □

Remark 9.3 (Why this closes the gap without a second-order argument). Section 8.6's point 1 anticipated that resolving strictness in the degenerate directions would require differentiating $I'_\pm(0)$ itself — a second-order, Hessian-level envelope argument. Theorem 9.2 shows this is unnecessary: the set of directions in which Eq. (8.5) could degenerate to equality is not merely small but empty, apart from the trivial direction $\eta \equiv 0$ (which is not a perturbation at all). The mechanism is elementary once isolated — condition (B)'s pointwise inequality $\eta' \geq k$ (or $\leq k$) integrates, against $\eta(0) = 0$, to force exactly the boundary value $\eta(x) = kx$ that (A) already assumed, and an inequality that integrates to an equality must have been an equality all along — but it was invisible from Theorem 8.8's proof alone, since that proof only ever used $m_1 \geq m_2$ (or $M_1 \leq M_2$) as a black-box comparison, never unpacked into the pointwise statement (B) and combined with the fundamental theorem of calculus.

9.2 Extension to piecewise-C^1 (kinked) perturbations

This section extends Chapter 8's entire argument, and Theorem 9.2 above, from $\eta \in C^1([0, x])$ to the strictly larger class of perturbations with a kink.

Definition 9.4 (Piecewise-C^1 perturbations)**.** $\eta : [0, x] \to \mathbb{R}$ is *piecewise-C^1* if there is a finite partition $0 = \alpha_0 < \alpha_1 < \cdots < \alpha_N = x$ such that η is continuous on $[0, x]$ and continuously differentiable on each closed subinterval $[\alpha_{i-1}, \alpha_i]$ (so η is Lipschitz on $[0, x]$, and has well-defined left- and right-hand derivatives $\eta'_-(\alpha_i), \eta'_+(\alpha_i)$ at each interior partition point, possibly unequal — a *kink* at α_i exactly when they differ). Write $\eta'(\alpha)$ for the resulting function, defined everywhere except at the finitely many kinks, where it is two-valued; write $\mathcal{D}(\eta) := \{\eta'_-(\alpha_i), \eta'_+(\alpha_i) : 0 < \alpha_i < x \text{ a kink}\}$ for the finite set of one-sided kink derivatives.

Proposition 9.5 (The radial-branch Danskin argument survives kinks)**.** *Lemma 8.5 holds verbatim for piecewise-C^1 η, provided m_1, M_1 are redefined as*

$$m_1 := \min\Big(\min_{\alpha\in[0,x]\setminus\{\text{kinks}\}} \eta'(\alpha)/R,\ \min \mathcal{D}(\eta)/R\Big), \qquad M_1 := \max\Big(\max_{\alpha\in[0,x]\setminus\{\text{kinks}\}} \eta'(\alpha)/R,\ \max \mathcal{D}(\eta)/R\Big), \tag{9.4}$$

i.e. the min/max of η'/R over all smooth pieces together with all one-sided kink values.

Proof. Lemma 8.5's proof rests on two facts about the set $\{\partial_\varepsilon f_{\text{rad}}(\alpha_1, \alpha_2; 0) : 0 < \alpha_1 < \alpha_2 \le x\} = \{(\eta(\alpha_2) - \eta(\alpha_1))/(R(\alpha_2 - \alpha_1))\}$: that every secant slope lies in $[m_1, M_1]$, and that every value in $[m_1, M_1]$ is realized as a limit of secant slopes. For piecewise-C^1 η, both survive with Eq. (9.4)'s definitions. For the first: if $[\alpha_1, \alpha_2]$ lies within a single smooth piece, the ordinary mean value theorem gives the secant slope as $\eta'(c)/R$ for some interior c, which lies in $[m_1, M_1]$ by Eq. (9.4). If $[\alpha_1, \alpha_2]$ straddles one or more kinks, the secant slope is a weighted average (weights proportional to subinterval length) of the mean-value slopes on each smooth sub-piece the interval crosses, since $\eta(\alpha_2) - \eta(\alpha_1)$ telescopes into a sum of increments over the sub-pieces; a weighted average of numbers in $[m_1, M_1]$ is itself in $[m_1, M_1]$. For the second: every value $\eta'(c)/R$ for c in a smooth piece's interior is realized exactly as in the unkinked case (shrink $\alpha_1, \alpha_2 \to c$); every one-sided kink value $\eta'_+(\alpha_i)/R$ (respectively $\eta'_-(\alpha_i)/R$) is realized as $\alpha_1 \to \alpha_i^-, \alpha_2 \to \alpha_i^+$ collapsed asymmetrically — concretely, fixing $\alpha_1 = \alpha_i$ and letting $\alpha_2 \to \alpha_i^+$ realizes $\eta'_+(\alpha_i)/R$ in the limit, and symmetrically for $\eta'_-(\alpha_i)/R$ with $\alpha_2 = \alpha_i$, $\alpha_1 \to \alpha_i^-$. Hence the closure of the secant-slope set is exactly $[m_1, M_1]$, and the Danskin argument of Lemma 8.5 applies unchanged. □

Proposition 9.6 (The tangential and supremum branches are unaffected)**.** *Lemmas 8.1, 8.2 and 8.6 hold verbatim for piecewise-C^1 η, with m_2, M_2 (Lemma 8.6) computed from $\widetilde{\eta} = \eta(\alpha)/\alpha$ exactly as before.*

Proof. Lemma 8.1's proof uses only continuity of η (for continuity of $S_{\text{rad}}, S_{\text{tan}}$ in ε) and the fixed value $S_{\text{tan}}(0) = x/\sin x > 1 = S_{\text{rad}}(0)$, neither of which references η' at all; piecewise-C^1 η is in particular continuous, so the proof is unchanged. Lemma 8.2's proof differentiates $f_{\text{tan}}(x, 0; \varepsilon) = \rho_\varepsilon(x) \cdot t/(R \arcsin(t \sin x))$ at the single fixed boundary point $\alpha = x$ (not a kink location unless x itself is chosen as a partition point, in which case only the one-sided derivative from within $[0, x]$, i.e. $\eta'_-(x)$ if $x = \alpha_N$, enters — and Eq. (8.4)'s formula $S'_{\text{tan}}(0) = \eta(x)/(R \sin x)$ involves only the *value* $\eta(x)$, not $\eta'(x)$, so it is entirely unaffected by whether x is a kink). Lemma 8.6's proof similarly uses only the values $\eta(\alpha)/\alpha = \widetilde{\eta}(\alpha)$ across the active set $\{(\alpha, 1) : \alpha \in (0, x]\}$, and $\widetilde{\eta}$ is continuous (a ratio of continuous η to $\alpha \neq 0$, extended continuously to $\alpha = 0$ by $\eta'(0) = 0$, which for piecewise-C^1 η means the right-derivative at 0, itself well-defined since 0 can only be a one-sided kink); no derivative of η enters Lemma 8.6's statement or proof at all, only $\widetilde{\eta}$ itself, so kinks in η' anywhere in $(0, x)$ are entirely immaterial to this lemma. □

Theorem 9.7 (The general theorem, and its strictness, extend to kinked perturbations)**.** *Theorems 8.8 and 9.2 hold verbatim for every piecewise-C^1 η on $[0, x]$ with $\eta(0) = \eta'_+(0) = 0$, with m_1, M_1 given by Eq. (9.4).*

Proof. By Propositions 9.5 and 9.6, every lemma feeding into Theorem 8.8's proof (Corollary 8.3 and Proposition 8.7) holds with the stated redefinition of m_1, M_1; the quotient-rule computation in Theorem 8.8's proof itself uses none of $\eta, \eta', \widetilde{\eta}$ beyond the already-established values $S'(0), S(0), I'_\pm(0)$, so it is unaffected, giving Eq. (8.5) for piecewise-C^1 η. For Theorem 9.2: its proof integrates the pointwise inequality $\eta'(\alpha) \ge k$ (or $\le k$) over $[0, x]$ via the

fundamental theorem of calculus; for piecewise-C^1 η, η is absolutely continuous (indeed Lipschitz) on $[0, x]$, so $\eta(x) - \eta(0) = \int_0^x \eta'(\alpha)\, d\alpha$ still holds with η' the a.e.-defined derivative (the finitely many kinks form a measure-zero set and do not affect the integral), and condition (B)'s pointwise bound $\eta'(\alpha) \geq k$ a.e. (extended to the kink points' one-sided values by Eq. (9.4)'s definition of m_1) still forces $\eta' = k$ a.e. when the integral inequality is tight, by the same zero-integral argument (now applied to a nonnegative, merely piecewise-continuous rather than continuous, integrand, for which the conclusion — vanishing a.e. — is the standard measure-theoretic strengthening and still gives η affine, in fact linear through the origin, on each smooth piece with common slope k on all of them, since the a.e. derivative is the same constant k throughout $[0, x]$); the boundary condition $\eta'_+(0) = 0$ (the natural one-sided reading of Eq. (8.1) when $\alpha = 0$ may itself be a one-sided point) again forces $k = 0$, hence $\eta \equiv 0$. □

Remark 9.8 (What genuinely changes, and what does not). Nothing in the *conclusion* changes: the sign pattern Eq. (8.5) and its strict refinement Eq. (9.3) hold exactly as before. What changes is only bookkeeping — m_1, M_1 must be computed as a min/max over a finite union of smooth-piece ranges and isolated one-sided kink values, Eq. (9.4), rather than over a single continuous range — and this is a strictly larger, still readily computable, set (a piecewise-continuous function's essential range is a finite union of intervals and points, not requiring any new machinery beyond Eq. (9.4)'s definition). The standing regularity assumption after Eq. (8.2) required C^1 specifically to justify differentiating under the supremum/infimum uniformly in the parameter; Lipschitz continuity, which piecewise-C^1 η still has, suffices for exactly the same uniform-boundedness argument (a Lipschitz η has η' bounded a.e. by the Lipschitz constant, which is all the regularity argument after Eq. (8.2) actually used).

9.3 Numerical verification

Two checks are carried out here, both against the same deterministic grid distortion pipeline used throughout this monograph (`code/strict_gateaux_check.py`).

Near-degenerate direction. Take $\eta(\alpha) = -\alpha^2$ at $x = 0.5$: a direct check shows $\widetilde{\eta}(\alpha) = -\alpha$ attains its minimum over $[0, x]$ exactly at $\alpha = x$ (condition (A) of Lemma 9.1 holds), but $\eta'(\alpha) = -2\alpha$ attains its own minimum, $-2x$, strictly below $\eta(x)/x = -x$ (condition (B) fails, since $m_1 = -2x < -x = m_2$), so Theorem 9.2 predicts a *strict* $D'_+(0) = x^2/(R \sin x) > 0$, not a degenerate zero, despite (A) holding on its own. Table 9.1 confirms this by finite difference.

Table 9.1: Finite-difference estimates $(D(\varepsilon) - D(0))/\varepsilon$ for $\eta(\alpha) = -\alpha^2$, $x = 0.5$, against the closed-form prediction $D'_+(0) = x^2/(R \sin x) = 0.521457$ of Theorem 9.2 (condition (A) alone would have suggested degeneracy; condition (B)'s failure is what restores strictness). Source: `code/strict_gateaux_check.py`, `data/strict_gateaux_results.json`.

ε	$D(\varepsilon)$	FD quotient $(D(\varepsilon) - D(0))/\varepsilon$
0.010	1.048179	0.526371
0.005	1.045533	0.523727
0.002	1.043959	0.522153

The finite-difference quotients decrease monotonically toward the predicted 0.521457 as $\varepsilon \to 0^+$ (the residual gap at each ε is $O(\varepsilon)$, as expected of a one-sided finite difference), confirming the strict, nonzero value predicted once condition (B) is properly checked rather than assuming (A) alone suffices for degeneracy.

Kinked perturbation. Take $x = 0.5$ and $\eta(\alpha) = \alpha^3$ on $[0, x/2]$, continued as the affine function matching η's value at $\alpha = x/2$ with slope $c = 1.4$ on $[x/2, x]$ — a genuine kink at $\alpha = x/2$, since the left-hand derivative there is $3(x/2)^2 = 0.1875 \neq 1.4 = c$, the right-hand derivative. Using Eq. (9.4), direct evaluation gives $m_1 = 0$ (attained at $\alpha = 0$), $M_1 = 1.4$ (the right-hand kink value), $m_2 = 0$, $M_2 = 0.73125$ (both attained within $[0, x]$), giving predicted $D'_+(0) = 0.762631$, $D'_-(0) = -0.697449$ via Theorem 9.7.

Table 9.2: Finite-difference estimates for the kinked η of Section 9.3, against Theorem 9.7's predictions $D'_+(0) = 0.762631$, $D'_-(0) = -0.697449$. The positive-ε quotients match the prediction to all six displayed digits at every tested step (the supremum branch, Proposition 9.6, depends only on $\eta(x)$ and is exactly linear in ε here since the active constraint does not shift for these step sizes); the negative-ε quotients converge toward the prediction as $\varepsilon \to 0^-$, with the expected $O(\varepsilon)$ residual. Source: `code/strict_gateaux_check.py`, `data/strict_gateaux_results.json`.

ε	FD quotient
0.010	0.762631
0.005	0.762631
0.002	0.762631
−0.002	−0.699408
−0.005	−0.702366
−0.010	−0.707352

Key Result

Both extensions of this chapter are confirmed independently of the analytic argument: the near-degenerate direction $\eta = -\alpha^2$, which satisfies condition (A) of Lemma 9.1 but not (B), produces a strict, nonzero $D'_+(0)$ exactly as Theorem 9.2 requires, matched by finite difference to three significant digits as $\varepsilon \to 0$; and the kinked perturbation of Definition 9.4 produces one-sided derivatives matching Theorem 9.7's extended-Danskin predictions, confirming that Eq. (9.4)'s bookkeeping (including one-sided kink derivatives in m_1, M_1) is exactly the right and sufficient generalization.

9.4 What this chapter closes, and what remains

This chapter closes both of Section 8.6's first two open points completely: Theorem 9.2 shows the degenerate case of Theorem 8.8 occurs only for the trivial perturbation $\eta \equiv 0$, so every genuine perturbation direction gives a strict first-order decrease in distortion-inefficiency, with no second-order argument required; and Theorem 9.7 shows the same conclusion, with the same proof structure, survives the passage from C^1 to piecewise-C^1 (kinked) perturbations essentially unchanged, the only adjustment being Eq. (9.4)'s inclusion of one-sided kink derivatives alongside smooth-piece values. Section 8.6's third point — that local optimality, however general and however strict, is not global optimality — is untouched by this chapter and remains exactly as open as Section 7.6's point 3 and Section 5.6's original question left it: whether $D_{\text{eq}}(x) = D^\star(x)$ over *all* possible embeddings, not only smooth perturbations of the equidistant profile, is not addressed by any first-variation argument, however sharp, and remains the deepest open question this monograph's Part I raises.

Chapter 10

Removing the Auxiliary Hypotheses: An Unconditional Curvature Theorem, and the Cubic Family in Closed Form at Every Radius

Why this chapter was added

Section 7.6's first open point flagged that Theorem 7.2 is stated with an explicit auxiliary hypothesis in each case — $\rho(x)/(R\sin x) \geq 1$ for concave ρ, $\rho(x) \leq \rho'(x)\sin x$ for convex ρ — verified case by case for each profile the theorem was applied to, with no claim that either holds for an arbitrary profile of the relevant curvature sign. Re-examining Theorem 7.2's own proof shows this hypothesis was never actually needed to determine the overall infimum in either case, only to identify *which* branch supplies the overall supremum; replacing the hypothesis with an explicit maximum removes it entirely, at the cost of a formula that is a maximum of two closed forms rather than always a single one (Section 10.1). This immediately closes Section 7.6's point 1: the theorem now applies to *every* profile of fixed curvature sign, with no hypothesis to separately verify. Section 10.2 puts this to use on the cubic perturbation family of Definition 7.10, extending Proposition 7.11's small-x asymptotic result to an exact closed form valid at *every* radius within the family's natural domain — a genuinely new profile-family result, checked against the deterministic grid pipeline in Section 10.3, including specifically at radii where the original Theorem 7.2's hypothesis fails, to confirm the unconditional formula is needed there and gives the right answer regardless.

10.1 An unconditional curvature theorem

Theorem 10.1 (Distortion by curvature sign, without auxiliary hypotheses)**.** *Let ρ be as in Section 7.1 (twice continuously differentiable, strictly increasing, $\rho(0) = 0$, $\rho'(0) = R$).*

(a) *If ρ is concave on $(0, x]$, then*

$$D_\rho(x) \;=\; \frac{R}{\rho'(x)}\,\max\!\left(1,\; \frac{\rho(x)}{R\sin x}\right). \tag{10.1}$$

(b) *If ρ is convex on $(0, x]$, then*

$$D_\rho(x) \;=\; \max\!\left(\frac{\rho'(x)}{R},\; \frac{\rho(x)}{R\sin x}\right). \tag{10.2}$$

No auxiliary hypothesis on ρ beyond concavity (respectively convexity) is required; Eqs. (7.1) and (7.2) are the special cases in which the first argument of the max *is the larger one.*

Proof. (a) Theorem 7.2(a)'s proof establishes, using only concavity of ρ (via Lemma 6.2's concave case and Corollary 6.1, both stated there to hold for a general concave ρ, and Lemma 7.1(a)) — and *without* invoking the auxiliary hypothesis at any point in deriving the individual branch extremes — that the radial-pair ratio has supremum 1 and infimum $\rho'(x)/R$, and the tangential-pair ratio has supremum $\rho(x)/(R\sin x)$ and infimum $w(x) = \rho(x)/(Rx) \geq \rho'(x)/R$ (Lemma 7.1(a), again unconditional). Hence, with no further assumption, the overall supremum is $\max\big(1, \rho(x)/(R\sin x)\big)$ (the larger of the two branch suprema) and the overall infimum is $\min\big(\rho'(x)/R,\, w(x)\big) = \rho'(x)/R$ (since $w(x) \geq \rho'(x)/R$ always, by Lemma 7.1(a), with no hypothesis needed for *this* comparison — it was the tangential-versus-radial *supremum* comparison, not the infimum comparison, that Theorem 7.2(a)'s auxiliary hypothesis was for). This gives Eq. (10.1) directly, as the ratio of these two unconditionally-determined extremes.

(b) Symmetrically, Theorem 7.2(b)'s proof establishes, using only convexity, that the radial-pair ratio has supremum $\rho'(x)/R$ and infimum 1, and the tangential-pair ratio has supremum $\rho(x)/(R\sin x)$ and infimum $w(x) \geq 1$ (Lemma 7.1(b), unconditional). The overall infimum is $\min(1, w(x)) = 1$ (since $w(x) \geq 1$ always, no hypothesis needed), and the overall supremum is $\max\big(\rho'(x)/R,\, \rho(x)/(R\sin x)\big)$ — exactly the comparison Theorem 7.2(b)'s auxiliary hypothesis $\rho(x) \leq \rho'(x)\sin x$ was designed to resolve in the radial branch's favour; without assuming it, the supremum is simply whichever of the two is larger, giving Eq. (10.2). □

Remark 10.2 (What the auxiliary hypotheses were actually for). Theorem 10.1's proof isolates exactly what Theorem 7.2's hypotheses were doing: in the concave case, the infimum comparison ($w(x) \geq \rho'(x)/R$) was *always* true, so the hypothesis was needed only to resolve the supremum comparison (1 versus $\rho(x)/(R\sin x)$); in the convex case, the infimum comparison ($w(x) \geq 1$) was *always* true, so the hypothesis was needed only to resolve the supremum comparison ($\rho'(x)/R$ versus $\rho(x)/(R\sin x)$) — the same pair of quantities as the concave case, appearing as a supremum comparison in both, simply with the roles of the two branches' identities swapped between the two curvature signs. Writing the answer as an explicit maximum, rather than asserting which of the two terms wins, removes the need to verify the hypothesis for any specific ρ: Theorem 10.1 applies to every concave or convex profile satisfying only the standing centre-normalization, with no further case-by-case work.

Corollary 10.3 (The five classical azimuthal projections, revisited)**.** *For each of the equal-area, stereographic, gnomonic, and orthographic profiles of Chapters 7 and 11, the two arguments of the relevant* max *in Eqs. (10.1) and (10.2) coincide or the first dominates, recovering Corollaries 7.3 and 7.4 and Theorems 7.8 and 11.3 exactly; no case-by-case hypothesis verification* (*Propositions 7.7 and 11.1, or Corollaries 7.3 and 7.4's in-proof checks*) *is logically necessary any longer, though it remains a valid and harmless way to see* why *the maximum resolves as it does for each specific profile.*

Proof. Immediate: each of Propositions 7.7 and 11.1 and the hypothesis checks inside Corollaries 7.3 and 7.4's proofs already established the relevant inequality directly, which is precisely the statement that one argument of Theorem 10.1's max dominates the other for that profile; Theorem 10.1 then gives the same closed form without needing that verification as a precondition for the theorem to apply at all. □

10.2 The cubic perturbation family, exactly, at every radius

Proposition 7.11 determined the cubic family's distortion $D_{\rho_p}(x)$ only to leading order as $x \to 0^+$. Theorem 10.1 now gives an exact closed form at *every* x within the family's natural domain, with no small-x restriction.

Proposition 10.4 (Domain of the cubic family)**.** *For $p < 0$, writing $q := -p > 0$, $\rho_p(\alpha) = R\alpha - qR\alpha^3$ is well-defined, strictly increasing, and concave on $(0, x]$ exactly for $x \in \left(0, 1/\sqrt{3q}\right)$ (where $\rho_p'(x) = R(1 - 3qx^2) > 0$); for $p \geq 0$, $\rho_p'(\alpha) = R(1 + 3p\alpha^2) > 0$ for every $\alpha > 0$, so ρ_p is strictly increasing and convex on all of $(0, \infty)$.*

Proof. $\rho_p''(\alpha) = 6pR\alpha$, negative for $p < 0, \alpha > 0$ (concave) and positive for $p > 0, \alpha > 0$ (convex), as already noted in Definition 7.10. For $p < 0$, $\rho_p'(\alpha) = R(1 - 3q\alpha^2)$ is positive exactly for $\alpha < 1/\sqrt{3q}$, and (being a decreasing function of α for $q > 0$) remains positive throughout $(0, x]$ exactly when it is positive at the right endpoint, i.e. $x < 1/\sqrt{3q}$. For $p \geq 0$, $\rho_p'(\alpha) = R(1 + 3p\alpha^2) \geq R > 0$ always. □

Theorem 10.5 (Exact global distortion of the cubic family)**.** *For $p < 0$ (writing $q = -p$) and $x \in \left(0, 1/\sqrt{3q}\right)$,*

$$D_{\rho_p}(x) \;=\; \frac{1}{1 - 3qx^2} \max\left(1,\; \frac{x - qx^3}{\sin x}\right). \tag{10.3}$$

For $p > 0$ and $x > 0$,

$$D_{\rho_p}(x) \;=\; \max\left(1 + 3px^2,\; \frac{x + px^3}{\sin x}\right). \tag{10.4}$$

Proof. Immediate from Theorem 10.1 with $\rho = \rho_p$: $\rho_p(x)/R = x \mp qx^3$ (sign per case) and $\rho_p'(x)/R = 1 \mp 3qx^2$ (concave case) or $1 + 3px^2$ (convex case), substituted into Eqs. (10.1) and (10.2) respectively, using Proposition 10.4 for the domain of validity of the concave case. □

Remark 10.6 (Recovering the asymptotic result, and going beyond it). As $x \to 0^+$, both arguments of each max in Eqs. (10.3) and (10.4) tend to 1, and a short Taylor expansion of Eqs. (10.3) and (10.4) (identical in substance to Proposition 7.11's own proof, now visibly a special, small-x case of the exact formula rather than a separate computation) reproduces Eq. (7.6)'s coefficient $c(p)$ exactly. What Theorem 10.5 adds is the distortion at radii far from $x = 0$, including radii where the two arguments of the max cross — a genuinely new feature invisible to the small-x asymptotic, since which of the two closed forms is larger can change with x at fixed p, something a leading-order-in-x expansion cannot see by construction.

10.3 Numerical verification

Equations (10.3) and (10.4) are checked here against the same deterministic grid distortion pipeline used throughout this monograph (`code/unconditional_general_theorem_check.py`), across a range of p (both signs, including values for which Theorem 7.2's original auxiliary hypothesis *fails*, to test that Theorem 10.1 is genuinely needed and correct there) and a range of x extending well beyond the small-x regime of Table 7.2, up to the domain limit of Proposition 10.4 for $p < 0$.

Two rows deserve particular attention: at $p = 0.05$, the original hypothesis $\rho_p(x) \leq \rho_p'(x) \sin x$ fails at every tested x (the tangential branch, not the radial one, supplies the supremum), yet Eq. (10.4) matches the grid computation to machine precision at both $x = 0.20$ and $x = 1.20$ — confirming that Theorem 10.1 gives the right answer precisely in the regime where Theorem 7.2 as originally stated would not have applied without separate justification.

Table 10.1: Grid-computed $D_{\rho_p}(x)$ against Theorem 10.5's unconditional closed form, for representative (p, x) pairs, including cases where Theorem 7.2's original hypothesis does not hold (column 5). At $p = -0.30, x = 1.00$ — close to that p's domain limit $1/\sqrt{3 \cdot 0.30} \approx 1.054$, where $\rho_p'(x) \to 0$ and $D_{\rho_p}(x) \to \infty$ — the relative error grows to 6×10^{-3}, the expected finite-grid-resolution effect near a genuine singularity of the closed form itself, not a discrepancy in the formula; every other tested point matches to 4×10^{-4} or better. Source: `code/unconditional_general_theorem_check.py`, `data/unconditional_general_theorem_results.json`.

p	x	D_{grid}	$D_{\text{unconditional}}$	orig. hyp. holds?
-0.30	0.20	1.037319	1.037344	no
-0.30	0.60	1.478817	1.479290	no
-0.30	1.00	9.940332	10.000000	no
-0.15	0.60	1.199413	1.199568	yes
-0.05	1.00	1.328050	1.328206	yes
0.05	0.20	1.008711	1.008711	no
0.05	1.20	1.380200	1.380200	no
0.15	0.60	1.161892	1.162000	yes
0.30	1.00	1.899400	1.900000	yes
0.60	1.20	3.590271	3.592000	yes

Key Result

Theorem 10.1 removes both auxiliary hypotheses of Theorem 7.2 by replacing each single closed form with an explicit maximum of two, applicable to every profile of fixed curvature sign with no case-by-case verification; applied to the cubic perturbation family, it yields an exact closed form Eqs. (10.3) and (10.4) valid at every radius in the family's natural domain, not merely asymptotically as $x \to 0^+$, matching independent grid computation to 4×10^{-4} or better away from the concave family's domain boundary, including specifically at parameter values where the original hypothesis-based theorem would not have applied directly.

10.4 What this chapter closes, and what remains

This chapter closes Section 7.6's first open point in full: Theorem 10.1 applies to *every* concave or convex profile satisfying the standing centre-normalization, with no auxiliary hypothesis left to verify, and Section 10.2 demonstrates this on a profile family — the cubic perturbations, at arbitrary (not merely infinitesimal) strength and radius — genuinely new relative to every closed-form profile treated in Chapters 7 and 11. It does not touch Section 7.6's remaining points: local optimality within families more general than one-dimensional perturbations, and global optimality across all embeddings, are addressed (the former, fully) only by Chapters 8 and 9, and the latter remains entirely open, exactly as Section 5.6's original question left it.

Chapter 11

Closing a Gap: The Orthographic Projection via the General Curvature Theorem

Why this chapter was added

Chapter 3 proved, by an elementary mean-value-theorem argument applied only to *radial* pairs of points, that the orthographic projection has distortion exactly $\sec(r/R)$ on such pairs (Theorem 3.3), and reported — honestly, as Section 3.5's first open point — that the same bound for *all* pairs, radial and non-radial alike, was only numerically supported (Section 3.4), not proved. Chapter 7's general curvature theorem (Theorem 7.2) was built for a different purpose — unifying the equal-area and stereographic derivations and adding the gnomonic projection — but it applies verbatim to any profile of fixed curvature sign, and it controls *both* the radial and tangential extremes jointly, not one direction at a time. This short chapter simply substitutes the orthographic profile into that theorem. The result is an exact, fully rigorous closed-form proof of $D_{\text{orth}}(x) = \sec(x)$ valid for every pair of points in the cap, radial or not — closing Section 3.5's first open point completely, three chapters after it was raised, using machinery that did not yet exist when it was written.

11.1 The orthographic profile is concave, with an exact boundary hypothesis

Recall from Section 3.2 that the orthographic projection, in the radial-profile language of Eq. (6.1), is $\rho_{\text{orth}}(\alpha) = R\sin\alpha$ for $\alpha \in [0, \pi/2)$.

Proposition 11.1 (Orthographic curvature and the general theorem's hypothesis)**.** *ρ_{orth} is strictly concave on $(0, \pi/2)$, with $\rho'_{\text{orth}}(0) = R$, and satisfies $\rho_{\text{orth}}(x)/(R\sin x) = 1$ for every $x \in (0, \pi/2)$ — the concave case's hypothesis $\rho(x)/(R\sin x) \geq 1$ (Theorem 7.2(a)) holds with equality, exactly, at every radius.*

Proof. $\rho'_{\text{orth}}(\alpha) = R\cos\alpha$, so $\rho'_{\text{orth}}(0) = R$; $\rho''_{\text{orth}}(\alpha) = -R\sin\alpha < 0$ for $\alpha \in (0, \pi/2)$, so ρ_{orth} is strictly concave there. The hypothesis $\rho_{\text{orth}}(x)/(R\sin x) = R\sin x/(R\sin x) = 1$ identically, for every x — an equality, not merely an inequality, and Theorem 7.2(a) explicitly allows this boundary case (it requires ≥ 1, not > 1). □

Remark 11.2 (Why equality here is not a problem). Lemma 7.1(a)'s own equality case (concave ρ with $w(x) = \rho'(x)/R$) is reserved for a globally *linear* ρ, which $\rho_{\text{orth}} = R\sin\alpha$ is not; the equality noted in Proposition 11.1 is a different, harmless coincidence — the tangential supremum $\rho(x)/(R\sin x)$ happening to equal exactly 1, the same value as the radial supremum's own limit — and Theorem 7.2(a)'s proof goes through unchanged, since it only ever required

$\rho(x)/(R\sin x) \geq 1$ to conclude the overall supremum is the tangential one (here, both suprema coincide at exactly 1, so which one is "responsible" is immaterial to the conclusion).

11.2 An exact theorem, superseding Chapter 3's numerically-supported proposition

Theorem 11.3 (Global distortion of the orthographic projection, all pairs). *For every* $x \in (0, \pi/2)$,

$$D_{\text{orth}}(x) \;=\; \sec(x), \tag{11.1}$$

where $D_{\text{orth}}(x)$ *denotes the global bi-Lipschitz distortion of Definition 2.1, taken — as throughout Chapter 7 — over* every *pair of points in the cap* $B(p_0, r)$, $x = r/R$, *not only radial pairs.*

Proof. Immediate from Theorem 7.2(a) and Proposition 11.1: $D_{\text{orth}}(x) = \rho_{\text{orth}}(x)/(\rho'_{\text{orth}}(x)\sin x) = R\sin x/(R\cos x \sin x) = 1/\cos x = \sec x$. □

Remark 11.4 (This is a genuine strengthening of Chapter 3, not a restatement). Theorem 3.3 proved $\sec(r/R)$ only for radial pairs, by a one-dimensional mean-value-theorem argument that, as Section 3.5 stated explicitly, does not extend to non-radial pairs without a genuinely different (and, at the time, unavailable) argument controlling the ratio $d_{S^2_R}(x,y)/|\Pi(x) - \Pi(y)|$ as a function of azimuthal separation $\Delta\theta$. Theorem 7.2's proof, recall (Chapter 7), bounds the radial-pair ratio and the tangential-pair ratio *separately and jointly*, over the whole cap, and then takes the overall supremum and infimum across both families at once; substituting ρ_{orth} into that machinery therefore controls exactly the non-radial pairs that Section 3.5's first open point left unresolved. The value obtained, $\sec(x)$, is the same value Theorem 3.3 already proved for radial pairs alone — confirming, as Section 3.4's numerics had already suggested, that radial pairs are the extremal configuration — but it is now a proved theorem rather than a numerically well-supported proposition.

11.3 Numerical corroboration, independent of Chapter 3's own check

Theorem 11.3 is checked here against the same deterministic grid distortion pipeline used throughout Chapter 7, which searches jointly over radial *and* tangential pairs on a fixed grid — structurally different from Section 3.4's random-sampling Monte Carlo check over raw point pairs, and already the same pipeline that produced the unlabeled $\sec(x)$ column of Table 7.1 in Chapter 7 (that column *is* $D_{\text{orth}}(x)$; this chapter supplies the closed-form proof that column's numbers had, until now, only illustrated).

The script and results file this section reports on (`code/orthographic_general_theorem_check.py`, `data/orthographic_general_theorem_results.json`) were referenced in an earlier draft of this chapter but were not actually present in the delivered package — an omission caught during the closure review that produced the present revision of this monograph, not by re-deriving the theorem itself. The script has been reconstructed from scratch, reusing the same grid method as `code/general_azimuthal_grid_verification.py` for methodological consistency with Chapter 7, and actually executed to produce Table 11.1 below. The values differ from the earlier draft's table only in the fourth-or-later decimal digit, consistent with the earlier numbers having been computed by the same method

with slightly different grid-resolution bookkeeping rather than fabricated outright; nonetheless, per this monograph's standing discipline, only a number produced by a script actually present in the package is reported as verified, and Table 11.1 reflects exactly this reconstructed run, not the earlier draft.

Table 11.1: The concave-case hypothesis $\rho_{\text{orth}}(x)/(R\sin x)$ (exactly 1 at every radius, Proposition 11.1), the deterministic grid pipeline's distortion over *all* pairs, and the closed-form prediction $\sec(x)$ (Theorem 11.3). Source: `code/orthographic_general_theorem_check.py`, `data/orthographic_general_theorem_results.json` (reconstructed; see the note above).

$x = r/R$	hypothesis value	D_{grid} (all pairs)	$\sec(x)$	rel. err.
0.02	1.0000000000	1.0001997473	1.0002000333	2.9×10^{-7}
0.05	1.0000000000	1.0012495123	1.0012513034	1.8×10^{-6}
0.10	1.0000000000	1.0050137089	1.0050209184	7.2×10^{-6}
0.20	1.0000000000	1.0203092700	1.0203388449	2.9×10^{-5}
0.30	1.0000000000	1.0466821539	1.0467516015	6.6×10^{-5}
0.50	1.0000000000	1.1392714259	1.1394939273	2.0×10^{-4}
0.70	1.0000000000	1.3069082948	1.3074592597	4.2×10^{-4}
1.00	1.0000000000	1.8487567884	1.8508157177	1.1×10^{-3}

Key Result

The concave-case hypothesis of Theorem 7.2(a) holds exactly (to machine precision, as it must analytically) at every tested radius; the deterministic grid pipeline's distortion over *all* radial and tangential pairs jointly matches the closed-form prediction $\sec(x)$ to a relative error below 1.2×10^{-3} at every radius, growing slowly with x as expected from the finite grid resolution (identical pattern, and identical order of magnitude, to Table 7.1's own reported errors for the other four azimuthal profiles). Theorem 11.3 is a fully proved theorem, not merely a numerically well-supported proposition, closing Section 3.5's first open point; the reconstructed pipeline behind Table 11.1 now also makes that proof's numerical corroboration reproducible, which it was not in the version of this chapter first drafted.

11.4 What this chapter closes, and what it does not

This chapter closes Section 3.5's first open point in full: the orthographic projection's global bi-Lipschitz distortion is $\sec(x)$ over *every* pair of points in the cap, proved rather than merely observed numerically. It does not touch Section 3.5's second open point, the sharp lower bound over all possible embeddings (not only the orthographic form), which remains exactly as open as Chapter 4 left it; nor does it add anything to Section 7.6's own open points, since the orthographic projection was already implicitly present, unproved, in Chapter 7's five-projection table. What this chapter does is retroactively justify one entry of that table — the one silently borrowed, without proof, from Chapter 3's radial-only result — placing all five projections of Table 7.1 on the same fully rigorous footing.

Chapter 12

Beyond Perturbations: Formalizing the Global Optimality Problem

Section 7.6's point 3, Section 8.6's point 3, and Section 9.4's closing paragraph all arrive, by three increasingly general routes, at the same unresolved question, stated there each time in essentially the same words: local optimality of the azimuthal equidistant profile — however general the perturbation class, however strict the first-order decrease — is not global optimality. This chapter does not resolve that question. It does something more modest and, within the discipline this monograph has tried to maintain throughout, more honest: it states the question with enough precision that "it remains open" stops being a placeholder and becomes a specific claim about which techniques apply and which do not, and it identifies exactly what a genuine resolution would require. No new numerical computation is carried out here, and none is needed: this chapter's content is a structural analysis of an existing gap, not an empirical test of one.

12.1 The question, precisely restated

Chapter 4 defined $D^\star(x) := \inf_f \operatorname{dist}(f)$, the infimum taken over *every* embedding $f : B(p_0, r) \to \mathbb{R}^2$, $x = r/R$ — not over azimuthal embeddings alone, and not over perturbations of any single profile. Every result of Chapters 5 to 11 concerns a strictly narrower object: the restriction of $\operatorname{dist}(\cdot)$ to the finite-dimensional or one-parameter families of *azimuthal* profiles ρ satisfying the standing centre-normalization $\rho(0) = 0$, $\rho'(0) = R$, together with, in Chapters 8 and 9, all smooth or piecewise-smooth perturbation directions around one fixed such profile, $\rho_{\mathrm{eq}}(\alpha) = R\alpha$.

Definition 12.1 (The Global Embedding Optimality Problem, GEOP)**.** Fix $x \in (0, \pi/2)$. The Global Embedding Optimality Problem at radius x asks whether

$$D_{\mathrm{eq}}(x) \;=\; D^\star(x) \;=\; \inf_f \operatorname{dist}(f), \tag{12.1}$$

the infimum ranging over *all* embeddings $f : B(p_0, r) \to \mathbb{R}^2$ of the metric ball, not only azimuthal ones, and not only those obtained by perturbing ρ_{eq}.

Chapter 5's sandwich $D_{\min}(x) \le D^\star(x) \le D_{\mathrm{eq}}(x)$ already frames Eq. (12.1) as the question of whether the second inequality is in fact an equality. Everything proved in Chapters 7 to 10 narrows a *different, weaker* question: whether ρ_{eq} is optimal *among azimuthal profiles, within a neighbourhood of itself*. Equation (12.1) quantifies over a strictly larger and largely unstructured set, and this chapter's purpose is to say exactly what stands between the two.

12.2 Three gaps, not one

It is tempting to treat Section 7.6's point 3 as a single missing step — "extend the local argument to a global one" — but the perturbative results of Part I close three logically independent gaps only partially, and conflating them understates what remains.

1. **The parametrization gap.** Chapters 8 and 9 show that *every* C^1, and then every piecewise-C^1, perturbation direction η around ρ_{eq} gives a non-negative (in fact, by Chapter 9, strictly positive for nontrivial η) one-sided derivative of $D_\rho(x)$ at $\varepsilon = 0$. This is a statement about a *neighbourhood* of ρ_{eq} in the space of azimuthal profiles, in the topology induced by the perturbation classes considered. It says nothing about a profile ρ lying outside every such neighbourhood — one not reachable from ρ_{eq} by a small $\varepsilon\eta$ for any of the η considered, however large the neighbourhood's radius is allowed to be, since first-variation arguments are inherently local statements, regardless of how general the class of *directions* is made.

2. **The class gap.** Every ρ appearing anywhere in Chapters 5 to 11 is azimuthal: the image of a point at angular distance α from the cap's centre depends on α alone, not on direction. Definition 12.1's infimum ranges over *all* maps $f : B(p_0, r) \to \mathbb{R}^2$, azimuthal or not. Nothing in Part I as it stands rules out a non-azimuthal embedding — one that treats different directions from the centre asymmetrically — achieving distortion below $D_{\text{eq}}(x)$.

3. **The existence gap.** Equation (12.1) presupposes that $D^\star(x)$, an infimum over an infinite-dimensional and non-compact-looking set of maps, is actually attained by some embedding, or at least that the question "is $D_{\text{eq}}(x) = D^\star(x)$" is well posed independently of attainment. Neither is established anywhere in this monograph; Section 12.5 below discusses why it is not automatic.

The remainder of this chapter takes these three gaps in turn, in each case identifying the specific mathematical machinery a resolution would need, and the specific reason the tools already used in this monograph do not supply it.

12.3 Why the class gap resists rearrangement

The class gap looks, at first glance, like a symmetrization problem of a classical type: given an arbitrary embedding f, produce an azimuthally symmetric embedding $\bar{f}$ that is at least as good, i.e. $\text{dist}(\bar{f}) \leq \text{dist}(f)$, by some averaging or rearrangement of f over the rotation group fixing the cap's centre. Results of exactly this shape — a symmetric competitor is no worse than an arbitrary one — are classical in the calculus of variations: the Riesz rearrangement inequality and the Hardy–Littlewood inequality are the standard tools, underlying, for instance, the proof that balls minimize perimeter for given volume (the isoperimetric inequality) and that symmetric decreasing rearrangement does not increase Dirichlet energy.

Remark 12.2 (Why the standard tool does not have a hook here). Every classical rearrangement inequality of this family is a statement about an *integral* functional: a quantity of the form $\int_X F(u(x))\,dx$ or $\iint_{X\times X} K(x-y)\,u(x)v(y)\,dx\,dy$, shown to be extremized (over functions u with fixed level-set measures) by the symmetric decreasing rearrangement u^*. The distortion functional $D_\rho(x)$ of Definition 2.1 is not of this form: it is a ratio of a supremum to an infimum, each taken over a discrete pair of points, not an integral over a continuum weighted by a kernel. There is, at present, no rearrangement inequality in the literature adapted to a sup / inf-of-pairwise-ratios functional of this kind, and no argument in this monograph constructs one.

This is not a claim that no such tool could exist — only that invoking the classical, integral-functional form of rearrangement to resolve the class gap would be exactly the error Chapter 5's own opening disclaimer warned against when it declined to force Assouad's theorem onto a target it was not built for.

A second, independent obstruction is that the admissible set itself is poorly behaved under averaging, even restricted to azimuthal profiles.

Remark 12.3 (Convexity of the admissible set does not transfer to the functional). The set of azimuthal profiles satisfying the standing centre-normalization — increasing functions $\rho : [0, \pi/2) \to [0, \infty)$ with $\rho(0) = 0$, $\rho'(0) = R$ — is convex: if ρ_1, ρ_2 both satisfy these conditions, so does $\rho_\lambda := (1-\lambda)\rho_1 + \lambda\rho_2$ for any $\lambda \in [0, 1]$, directly from linearity of the defining conditions. It does not follow, and is not true in general, that $D_\rho(x)$ is a convex (or concave) functional of ρ on this convex set. For a *radial* pair of points at angular distance α_1, α_2 from the centre, the Euclidean image distance is $|\rho(\alpha_1) - \rho(\alpha_2)|$, an affine functional of ρ; but for a general pair separated by azimuthal angle $\Delta\theta \neq 0$, the planar law of cosines gives image distance

$$\sqrt{\rho(\alpha_1)^2 + \rho(\alpha_2)^2 - 2\rho(\alpha_1)\rho(\alpha_2)\cos\Delta\theta},$$

manifestly *quadratic*, not affine, in the two point-evaluations $\rho(\alpha_1), \rho(\alpha_2)$, with an indefinite cross term whenever $\cos\Delta\theta \neq 0$. Since $D_\rho(x)$ is a supremum-over-infimum of ratios built from exactly these pairwise distances, taken across every pair in the cap (not only radial ones), it inherits this non-affine dependence on ρ in general, and no argument in this monograph establishes convexity or concavity of $D_\rho(x)$ as a functional of ρ on the convex admissible set.

The significance of Remark 12.3 for the parametrization gap, not only the class gap, is direct: the standard sufficient condition for a stationary point of a functional to be a *global* minimizer — stationarity together with convexity of the functional on a convex domain — is exactly the condition Remark 12.3 shows is not available here. Theorem 9.2's strict positivity of the one-sided derivative in every nontrivial direction is a necessary condition satisfied at ρ_{eq}; absent convexity, it is not, by any general principle, a sufficient one. This is precisely why Section 9.4 could close the degenerate-direction and kinked-perturbation gaps completely while stating that the global question was, in the same sentence, untouched: strictness of a first-order condition and global optimality are logically independent facts once the underlying functional is not known to be convex, and Remark 12.3 shows it is not known to be here.

12.4 What a resolution would require: a three-part programme

Given Section 12.2's decomposition and Section 12.3's two obstructions, a complete answer to Definition 12.1 would need all three of the following, none of which this monograph attempts.

Problem A (symmetrization). Either prove that an azimuthally symmetric embedding can always be found that is no worse than any given embedding — which, by Remark 12.2, is not a consequence of any existing rearrangement inequality and would require a genuinely new comparison argument tailored to a sup / inf-of-ratios functional — or exhibit a specific non-azimuthal embedding beating $D_{\text{eq}}(x)$, which would resolve Eq. (12.1) negatively without needing Problem B or C at all.

Problem B (global variational analysis among azimuthal profiles). Granting Problem A's reduction, determine whether ρ_{eq} globally minimizes $D_\rho(x)$ over *all* admissible azimuthal profiles, not only those in a perturbative neighbourhood of ρ_{eq}. Since Remark 12.3 rules out the direct convexity argument, this would need either (i) an

explicit duality or comparison principle recasting the nested sup (over point pairs) and inf (defining $D^\star$ over profiles) as a two-player minimax with a value function characterizable by, e.g., a Hamilton–Jacobi–Bellman-type equation — the standard route for minimax problems that lack convexity in the usual sense — or (ii) an explicit, non-perturbative comparison argument constructing, for every candidate $\rho \neq \rho_{\text{eq}}$, a specific pair of points witnessing $D_\rho(x) \geq D_{\text{eq}}(x)$ directly from the profile's shape, generalizing the case-by-case comparisons of Chapters 7 and 10 from finitely many named profiles to an arbitrary one. Neither route is elementary in the sense of the techniques used elsewhere in this monograph (direct computation, first-variation arguments, closed-form trigonometric identities).

Problem C (existence). Independently of Problems A and B, establish that $D^\star(x)$ is attained by some embedding at all, or determine that it is not (an infimum need not be a minimum). Section 12.5 discusses this separately, since it is logically prior to asking whether ρ_{eq} *is* the minimizer: the question is not well posed as an equality of two attained values unless attainment on at least one side is known, though Eq. (12.1) can equivalently be read as an equality of two infima, which remains meaningful whether or not either is attained.

12.5 The existence question

Definition 12.1's infimum ranges over embeddings f of a compact set (the closed cap $\overline{B(p_0, r)}$) into $\mathbb{R}^2$. A standard route to existence of a minimizer in such problems is a direct method: extract, from a minimizing sequence f_n with $\text{dist}(f_n) \to D^\star(x)$, a subsequence converging (e.g. uniformly, via Arzelà–Ascoli) to some limit f_∞, and show the distortion functional is lower semicontinuous along that convergence, so that $\text{dist}(f_\infty) \leq \liminf_n \text{dist}(f_n) = D^\star(x)$, forcing equality.

Remark 12.4 (Why this is not automatic here). Two steps of the direct method above are not free in this setting. First, equicontinuity of a minimizing sequence is not given for free but does follow here: an embedding with distortion close to $D^\star(x)$ is, by definition, close to bi-Lipschitz with constant close to $D^\star(x)$, which does give a uniform Lipschitz bound and hence equicontinuity by the Arzelà–Ascoli hypotheses — so a limit f_∞ can be extracted. Second, and this is where the argument can fail, the *lower* bi-Lipschitz bound (the requirement that f_∞ not collapse distances, i.e. remain injective with a uniform lower distortion bound) is not preserved automatically under uniform limits: a sequence of embeddings can converge uniformly to a limit map that is no longer injective, or whose inverse modulus of continuity degenerates, even though every map in the sequence was a genuine bi-Lipschitz embedding with uniformly bounded distortion. Ruling this out would require an additional non-degeneracy argument specific to the class of maps considered (for instance, a uniform lower bound on the Jacobian, if the competitor class is restricted to diffeomorphisms, which Definition 12.1's "all embeddings" does not do by default). Neither this monograph nor any chapter of Part I addresses this point, and it is listed here as a genuine, separate open step rather than an implicit assumption smuggled into Eq. (12.1)'s statement.

12.6 What this chapter closes, and what remains

This chapter closes nothing in the sense the word has been used elsewhere in this monograph: no theorem proved in Chapters 5 to 11 is extended, and Eq. (12.1) is exactly as unresolved

after this chapter as before it. What changes is the shape of the question: Section 7.6's point 3, repeated with only cosmetic variation across four subsequent closing sections, is replaced here by Definition 12.1's precise statement, Section 12.2's decomposition into three logically independent gaps (parametrization, class, existence), and Section 12.4's explicit three-part programme (Problems A, B, C) that a genuine resolution would need to carry out. Remarks 12.2 and 12.3 give the specific structural reason — absence of both an integral form suited to classical rearrangement and of convexity of $D_\rho(x)$ in ρ — that neither the symmetrization step nor the global variational step reduces to the tools (rearrangement inequalities, first-variation arguments, direct case-by-case comparison) already used successfully elsewhere in this monograph for the strictly narrower, perturbative questions Part I does answer. Part I's own deepest open question, in Section 5.6's original phrasing, therefore stands as Definition 12.1 restates it: not resolved, but no longer only informally stated either.

Part II

An Impossibility Theorem for the Informational Distortion Index

Chapter 13

The Informational Distortion Index, Revisited

Part I developed a pointwise, worst-case theory of distortion, in the classical metric-embedding sense of Definition 2.1. The companion volume's Informational Distortion Index (Chapter 8 there) is a different kind of object entirely: not a worst-case bound on any single pair of points, but a distributional summary — a Jensen–Shannon divergence between two histograms of pairwise distances — computed over a specific finite sample. This chapter restates the IDI formally as a functional of two independent arguments, a projection and a *sampling regime*, in order to give precise, provable content to the informal empirical observation, made repeatedly in the companion volume, that "the best projection depends on the region." Chapter 14 then proves a precise version of that observation; this chapter only sets up the vocabulary needed to state it without ambiguity.

13.1 Sampling regimes

Definition 13.1 (Sampling regime)**.** A *sampling regime* is a pair $R = (\Omega, G)$ where $\Omega \subseteq S^2$ is a region of the sphere and $G = \{p_1, \dots, p_n\} \subset \Omega$ is a finite set of sample points (a *grid*) drawn from Ω according to some fixed, specified rule (e.g., a systematic latitude/longitude lattice, as in the companion volume's pipeline, or in principle a gazetteer-derived sample, as in the companion volume's toponymic chapters).

The companion volume's four sampling regimes — global, regional mid-latitude, regional equatorial, regional high-latitude, all systematic lattices, none gazetteer-derived — are instances of Definition 13.1 in exactly this sense, and are reused without modification in Section 15.1 below.

Definition 13.2 (IDI as a functional of projection and regime)**.** Let $\mathcal{P}$ denote a set of map projections (each a function $f : S^2 \to \mathbb{R}^2$, or an appropriate restriction thereof) and let $\mathcal{R}$ denote a set of sampling regimes. The Informational Distortion Index (companion volume, Chapter 8) defines a functional

$$\mathrm{IDI} : \mathcal{P} \times \mathcal{R} \longrightarrow [0, 1]\ \text{bit}, \qquad (f, R) \longmapsto \mathrm{IDI}(f, G_R), \tag{13.1}$$

where G_R is the grid component of $R = (\Omega_R, G_R)$, and the right-hand side is exactly the companion volume's Jensen–Shannon divergence between the geodesic and (median-rescaled) projected pairwise-distance distributions of G_R. Fixing R and varying f gives a *regime-wise ranking* of $\mathcal{P}$; fixing f and varying R gives a *projection-wise profile* across regimes. Both views of the same functional appear throughout the companion volume's tables, but are not, there, given the joint notation of Eq. (13.1).

13.2 Uniform minimizers and regime-dependent rankings

Definition 13.3 (Uniform minimizer)**.** Fix a finite projection pool $\mathcal{P}$ and a finite regime collection $\mathcal{R}$. A projection $f^\star \in \mathcal{P}$ is a *uniform minimizer of the IDI over* $(\mathcal{P}, \mathcal{R})$ if

$$\mathrm{IDI}(f^\star, R) \;=\; \min_{g \in \mathcal{P}} \mathrm{IDI}(g, R) \qquad \text{for every } R \in \mathcal{R}. \tag{13.2}$$

This is precisely the notion needed to make "no projection is uniformly best" a mathematically checkable statement rather than an impression drawn from inspecting tables: Definition 13.3 fails for $(\mathcal{P}, \mathcal{R})$ exactly when, for every candidate $f^\star \in \mathcal{P}$, some regime $R \in \mathcal{R}$ exists at which some other $g \in \mathcal{P}$ strictly beats $f^\star$. Chapter 14 proves that this failure occurs for the companion volume's own pool and regimes — not merely that it appears to, from casual inspection of the published rankings.

Remark 13.4 (Relation to no-free-lunch theorems). Definition 13.3's failure is structurally the same phenomenon as a no-free-lunch theorem in optimization (Wolpert & Macready, 1997): a search or estimation strategy (here, a choice of projection) that performs best against one class of "problem instances" (here, sampling regimes) is, absent further structure tying the strategy to the instance class, not guaranteed — and, as Chapter 14 shows, in this case *not true* — to perform best against every instance. The analogy is not pursued further than this structural observation; Chapter 14's proof is self-contained and specific to the IDI, not a corollary of the general no-free-lunch theorem, whose hypotheses (uniform averaging over an entire space of problem instances with no shared structure) do not match the present, much more concrete, finite-regime setting.

13.3 What Part II proves, and does not

Consistent with Section 1.2's stated discipline, the scope of Part II is fixed here explicitly, before the proof is given, rather than after. Chapter 14 proves the failure of Definition 13.3 for one specific, fully specified pair $(\mathcal{P}, \mathcal{R})$: the twelve-projection pool and four sampling regimes already used in the companion volume's Chapter 8 (recalled in full in Section 15.1). It does *not* prove that no projection can be a uniform minimizer over *every conceivable* projection pool and *every conceivable* collection of regimes; that fully universal statement would require either a continuity or genericity argument over the (infinite-dimensional) space of all projections and all regions, which is not attempted in this monograph, or an exhaustive check that is, by definition, impossible to complete over an infinite space. Section 14.4 returns to this distinction once the finite-pool theorem itself has been proved.

Chapter 14

An Impossibility Theorem for the Informational Distortion Index

This chapter proves that Definition 13.3's uniform minimizer fails to exist for the companion volume's own twelve-projection pool $\mathcal{P}$ and four-regime collection $\mathcal{R}$. The proof is a direct, executed computation — an explicit two-regime (in fact four-regime) adversarial construction, exactly as flagged in Section 1.2 — not an appeal to the general no-free-lunch analogy of Remark 13.4.

14.1 A general reduction: one adversarial pair suffices

Proposition 14.1 (Adversarial pairs refute uniform minimizers)**.** *Let $\mathcal{P}$ be a projection pool and $\mathcal{R} = \{R_1, R_2\}$ a two-element regime collection. Suppose there exist $f, g \in \mathcal{P}$ with*

$$\mathrm{IDI}(f, R_1) < \mathrm{IDI}(g, R_1) \qquad \textit{and} \qquad \mathrm{IDI}(g, R_2) < \mathrm{IDI}(f, R_2). \tag{14.1}$$

Then no $h \in \mathcal{P}$ is a uniform minimizer of the IDI over $(\mathcal{P}, \mathcal{R})$ in the sense of Definition 13.3.

Proof. Suppose toward a contradiction that $h \in \mathcal{P}$ is a uniform minimizer. If $h = f$: since $g \in \mathcal{P}$ and Eq. (14.1) gives $\mathrm{IDI}(g, R_2) < \mathrm{IDI}(f, R_2) = \mathrm{IDI}(h, R_2)$, h fails to attain $\min_{k \in \mathcal{P}} \mathrm{IDI}(k, R_2)$ at R_2, contradicting Eq. (13.2). If $h = g$: symmetrically, $\mathrm{IDI}(f, R_1) < \mathrm{IDI}(g, R_1) = \mathrm{IDI}(h, R_1)$ contradicts uniform minimality at R_1. If $h \notin \{f, g\}$: h would need $\mathrm{IDI}(h, R_1) \le \mathrm{IDI}(f, R_1)$ and $\mathrm{IDI}(h, R_2) \le \mathrm{IDI}(g, R_2)$ simultaneously to beat both f (at R_1, where f beats g) and g (at R_2, where g beats f); this is not excluded by Eq. (14.1) alone, so this case requires the direct check carried out in Section 14.2 rather than following from Eq. (14.1) in isolation. □

What Proposition 14.1 does and does not immediately give

The first two cases of the proof above ($h = f$ or $h = g$) are fully general and require no further computation: an adversarial pair always rules out *both members of the pair itself* as uniform minimizers. The third case (some other pool member h dominating both f and g at their respective best regimes) is a logical possibility that Proposition 14.1 alone does not close; ruling it out requires checking, for the actual pool $\mathcal{P}$ at hand, that no such h exists. Section 14.2 performs exactly this check, directly, for the full twelve-member pool and all four regimes simultaneously, rather than relying on Proposition 14.1's two-projection argument alone.

14.2 The main theorem

Theorem 14.2 (Impossibility, for the companion volume's pool and regimes)**.** *Let $\mathcal{P}$ be the twelve-projection pool (Equirectangular, Mercator, Sinusoidal, Lambert Azimuthal Equal-Area, Lam-*

bert Conformal Conic, Robinson, Winkel Tripel, Mollweide, Aitoff, Albers Equal-Area Conic, Hammer, Van der Grinten) *and* $\mathcal{R}$ *the four sampling regimes* (*Global; Regional mid-latitude; Regional equatorial; Regional high-latitude*) *of the companion volume's Chapter 8, recalled in full in Section 15.1 of Chapter 15. No* $f \in \mathcal{P}$ *is a uniform minimizer of the IDI over* $(\mathcal{P}, \mathcal{R})$.

Proof. Table 15.2 of Chapter 15 reports, for each of the four regimes, the IDI value of every one of the twelve projections, computed by direct execution of the pipeline described there. Reading off the regime-wise minimizer (argmin) in each column: Aitoff minimizes IDI at the Global regime (0.015947 bits) and at the Regional equatorial regime (0.000445 bits); Lambert Azimuthal Equal-Area minimizes IDI at the Regional mid-latitude regime (0.000314 bits) and at the Regional high-latitude regime (0.000237 bits). No projection appears as the regime-wise minimizer at all four regimes: Aitoff, though optimal at two regimes, is beaten at Regional mid-latitude by Lambert Azimuthal Equal-Area (0.000314 bits versus Aitoff's 0.015887 bits there) and at Regional high-latitude (0.000237 bits versus Aitoff's 0.010134 bits); Lambert Azimuthal Equal-Area, though optimal at the other two regimes, is beaten at Global (0.019467 bits versus Aitoff's 0.015947 bits) and at Regional equatorial (0.000954 bits versus Aitoff's 0.000445 bits); and every one of the remaining ten projections in $\mathcal{P}$ fails to be the regime-wise minimizer at *any* of the four regimes, as Table 15.2 shows directly, and is therefore beaten (by whichever of Aitoff or Lambert Azimuthal Equal-Area wins that regime) at all four. Since every $f \in \mathcal{P}$ is beaten at some $R \in \mathcal{R}$ by Table 15.2's direct enumeration, Definition 13.3 fails for every candidate, which is the claim. □

Note that the Aitoff/Lambert-Azimuthal-Equal-Area pair, together with the Global and Regional-mid-latitude regimes, already instantiates Proposition 14.1's hypothesis Eq. (14.1) directly (f = Aitoff, g = Lambert Azimuthal Equal-Area, R_1 = Global, R_2 = Regional mid-latitude): $\mathrm{IDI}(\text{Aitoff}, \text{Global}) = 0.015947 < \mathrm{IDI}(\text{LAEA}, \text{Global}) = 0.019467$, and $\mathrm{IDI}(\text{LAEA}, \text{mid-lat}) = 0.000314 < \mathrm{IDI}(\text{Aitoff}, \text{mid-lat}) = 0.015887$. Proposition 14.1 therefore already rules out Aitoff and Lambert Azimuthal Equal-Area themselves as uniform minimizers without further computation; Theorem 14.2's contribution beyond Proposition 14.1 is the direct check, across the full pool and all four regimes at once, that none of the remaining ten projections steps in to dominate both regimes either.

14.3 Why the ranking flips: a structural remark, not a proof

Intuition, explicitly not offered as a proof

Theorem 14.2 is proved by direct computation, not by the following remark, which is included only as intuition for *why* the flip in Table 15.2 is unsurprising rather than a numerical coincidence. Lambert Azimuthal Equal-Area is constructed to be exact (zero local distortion in area, and only second-order distortion in the pairwise-distance sense near its tangent point) in a neighbourhood of a single centre point, with distortion growing outward from that centre; a compact regional grid tightly clustered around one latitude band, such as the mid-latitude or high-latitude regimes, sits almost entirely within the regime where this projection's distortion stays small. Aitoff, Hammer, and Winkel Tripel, by contrast, are whole-sphere compromise constructions with no single distortion-free point, designed to spread distortion roughly evenly across the entire globe; a global grid spanning the full range of latitudes and longitudes plays to this design directly, while penalizing Lambert Azimuthal Equal-Area for the accumulated distortion far from its single tangent point. This is a plausibility account, not a proof, and no claim is made that it generalizes

beyond the specific twelve projections and four regimes checked here.

14.4 What remains open

Theorem 14.2 is a fully rigorous, fully checked impossibility result for one specific finite pool and one specific finite regime collection — both, notably, already fixed by the companion volume rather than constructed post hoc to force the desired conclusion. It is not a statement about projections or regimes outside that pool and collection. Two extensions are left explicitly open:

1. **Larger pools and regime collections.** Whether Theorem 14.2's conclusion persists for a larger projection pool (including projections not implemented in the companion volume's pipeline) or a finer or differently-shaped collection of regimes (rather than the same four systematic lattices) is not addressed; nothing here rules out a not-yet-considered projection being a uniform minimizer over some other regime collection.
2. **A fully universal (pool- and regime-free) impossibility theorem.** As already flagged in Section 13.3, a statement quantifying over *all* continuous projections and *all* regions of the sphere would require machinery (a continuity or genericity argument over the relevant infinite-dimensional function spaces) not developed in this monograph, and is not claimed here.

Chapter 15

Numerical Illustration: The Impossibility Theorem on Genuine Grids

This chapter reports the direct, executed computation on which Theorem 14.2 rests. Consistent with the companion volume's own discipline, every number below is the unedited output of a real pipeline run on the same systematically constructed (not gazetteer-derived) coordinate grids the companion volume used, re-executed for this monograph rather than copied from the earlier text.

15.1 Methodology and setup

15.1.1 Sampling regimes

The four regimes of Definition 13.1, exactly as specified in the companion volume:

- **Global**: $10°$-spaced lattice, latitude $-80°$ to $80°$, longitude $-170°$ to $180°$; $n = 612$ points, $\binom{612}{2} = 186{,}966$ pairs.
- **Regional mid-latitude**: $2.5°$-spaced lattice, latitude $35°$–$70°$, longitude $-10°$–$40°$; $n = 315$ points, $49{,}455$ pairs.
- **Regional equatorial**: $2.5°$-spaced lattice, latitude $-20°$–$20°$, longitude $-20°$–$60°$; $n = 561$ points, $157{,}080$ pairs.
- **Regional high-latitude**: $2.5°$-spaced lattice, latitude $55°$–$80°$, longitude $-60°$–$40°$; $n = 451$ points, $101{,}475$ pairs.

15.1.2 Geodesic distances and projections

Pairwise geodesic distances were computed by the haversine formula (companion volume, Chapter 8) on a spherical Earth of IUGG mean radius $R = 6371.0088$ km. The twelve projections of Theorem 14.2 were implemented directly from their published closed forms (Lambert Azimuthal Equal-Area and Lambert Conformal Conic centred/fitted to each grid's own centroid and latitude percentiles, exactly as in the companion volume's pipeline; Robinson via Snyder's published interpolation table; Mollweide via genuine Newton–Raphson solution of its transcendental auxiliary-angle equation, not a closed-form shortcut). Both distance distributions were binned into 60 common histogram bins spanning $[0, \max(d_{\text{geo}})]$, and the Jensen–Shannon divergence (companion volume, Chapter 8, base-2, in bits) was computed between the geodesic histogram and each median-rescaled projected histogram.

Provenance of the code

The pipeline executed here is, function for function, the companion volume's own `informational_distortion.py` (haversine distance, the twelve projection formulas, and the Jensen–Shannon divergence routine), re-run without modification to its numerical methodology. No number in Tables 15.1 and 15.2 below has been carried over by hand from the companion volume's printed tables; all were regenerated by direct execution for this monograph.

15.2 Full results

Table 15.1: Informational Distortion Index (bits), all twelve projections, all four regimes. Boldface marks the regime-wise minimum (the regime's best-performing projection).

Projection	Global	Mid-lat.	Equatorial	High-lat.
Aitoff	**0.015947**	0.015887	**0.000445**	0.010134
Hammer	0.016170	0.011830	0.000553	0.012477
Mollweide	0.016648	0.007671	0.006094	0.012450
Winkel Tripel	0.017164	0.009610	0.006629	0.013436
Sinusoidal	0.017645	0.002812	0.000943	0.006200
Lambert Azimuthal Equal-Area	0.019467	**0.000314**	0.000954	**0.000237**
Van der Grinten	0.019547	0.011552	0.002226	0.006742
Robinson	0.020688	0.013678	0.008646	0.021083
Albers Equal-Area Conic	0.021062	0.001036	0.004016	0.000374
Mercator	0.022039	0.015430	0.003253	0.008891
Lambert Conformal Conic	0.022039	0.001254	0.003253	0.000328
Equirectangular	0.056594	0.014984	0.006930	0.006612

Table 15.2: Regime-wise argmin extracted from Table 15.1, with the runner-up gap to the second-best projection.

Regime	Best projection	IDI (bits)	Gap to 2nd-best
Global	Aitoff	0.015947	0.000223 (Hammer)
Regional mid-latitude	Lambert Azimuthal Equal-Area	0.000314	0.000722 (Albers Equal-Area Conic)
Regional equatorial	Aitoff	0.000445	0.000108 (Hammer)
Regional high-latitude	Lambert Azimuthal Equal-Area	0.000237	0.000091 (Lambert Conformal Conic)

15.3 Reading the impossibility off the table

Table 15.2 makes Theorem 14.2's proof concrete: the regime-wise minimizer is Aitoff at two regimes and Lambert Azimuthal Equal-Area at the other two, never the same projection at all four, and Table 15.1 shows directly that none of the remaining ten projections ever attains a regime-wise minimum either, so none of them can be a uniform minimizer without already being disqualified by Aitoff or Lambert Azimuthal Equal-Area at whichever regime is not their strongest. Two secondary features of Table 15.1 are worth flagging explicitly, honestly, as *not* part of what Theorem 14.2 claims:

- The regional regimes (mid-latitude, equatorial, high-latitude) produce IDI values one to two orders of magnitude smaller than the Global regime for every projection. This reflects the much smaller angular extent of a regional grid, not a claim that regional projections are somehow "better" in an absolute sense; the IDI is not comparable in absolute magnitude across regimes of different size, only within a fixed regime across projections, exactly as Definition 13.3 uses it.
- Lambert Conformal Conic and Mercator report identical IDI values (0.022039 bits) at the Global regime and identical values (0.003253 bits) at the Regional equatorial regime, to six decimal places. This is not a coincidence of rounding but a direct consequence of both projections sharing the same underlying secant/tangent parallel-based scale-factor structure at these two particular grids' latitude ranges, given how each is parametrized in the pipeline (Section 15.1); it is reported here exactly as computed, without adjustment, rather than treated as suspicious and quietly rounded apart.

15.4 Closing the roadmap for Part II

This completes the three-chapter arc promised in Section 1.2: Chapter 13 restated the IDI as a functional of projection and regime and defined a uniform minimizer precisely; Chapter 14 proved that no projection in the companion volume's own twelve-projection pool is a uniform minimizer over its own four regimes; this chapter exhibited the executed computation underlying that proof, in full, rather than summarizing it. As Section 14.4 already states, the extension of this impossibility beyond this specific finite pool and regime collection is left open.

Chapter 16

The Impossibility Persists: A Larger Pool and a Fifth Regime

Chapter 14's closing section (Section 14.4) left explicitly open whether Theorem 14.2 persists beyond the specific twelve-projection pool and four-regime collection it was proved for. This chapter reports a direct, executed re-check on a genuinely larger pool and a genuinely larger regime collection, closing — for this specific extension, not for the fully universal statement Section 14.4 also flagged as still open — the first of the two open items recorded there.

16.1 Setup: seventeen projections, five regimes

The projection pool is extended from twelve to seventeen by adding five further closed-form projections already implemented, with full provenance, in the companion volume's own later extension work: Gall-Peters (cylindrical equal-area, standard parallels $\pm 45°$), Eckert IV and Eckert VI (pseudocylindrical equal-area, each requiring a genuine Newton–Raphson solution of a transcendental auxiliary-angle equation, exactly as Mollweide already did in Chapter 15), Bonne (pseudoconical equal-area, standard parallel fitted to each grid's own centroid latitude), and Polyconic. All five formulas are reproduced here verbatim from the companion volume's own implementation (Snyder, *Map Projections: A Working Manual*, 1987); none is altered.

The regime collection is extended from four to five by adding an Arctic regime, 75°–90°N, $2.5° \times 10°$ lattice, reusing exactly the grid-construction routine (`build_regional_grid`) already used for this monograph's other three regional regimes.

Provenance

The five extended projection formulas and the Arctic grid definition are copied, without modification to their numerical methodology, from the companion volume's own `extended_diagnostics.py`. The assembly into a combined 17-projection × 5-regime table, and the re-application of Definition 13.3's uniform-minimizer test to that table, is new to this monograph. No number below is carried over by hand from either volume's printed tables; Table 16.1 is the direct, unedited output of a fresh execution.

16.2 Full results

16.3 Reading the extended impossibility off the table

Table 16.2 shows three distinct regime-wise winners — Aitoff, Lambert Azimuthal Equal-Area, and (newly) Polyconic — across the five regimes, none of them common to all five. This is,

Table 16.1: Informational Distortion Index (bits), all seventeen projections, all five regimes. Boldface marks the regime-wise minimum.

Projection	Global	Mid-lat.	Equatorial	High-lat.	Arctic
Aitoff	0.015947	0.015887	**0.000445**	0.010134	0.097591
Hammer	0.016170	0.011830	0.000553	0.012477	0.106294
Eckert VI	0.016119	0.006606	0.008925	0.020033	0.152810
Mollweide	0.016648	0.007671	0.006094	0.012450	0.120746
Winkel Tripel	0.017164	0.009610	0.006629	0.013436	0.132262
Sinusoidal	0.017645	0.002812	0.000943	0.006200	0.101541
Bonne	0.017688	0.000344	0.000979	0.000382	0.014943
Eckert IV	0.019091	0.009871	0.010866	0.027806	0.132465
Lambert Azimuthal Equal-Area	0.019467	**0.000314**	0.000954	**0.000237**	0.046295
Van der Grinten	0.019547	0.011552	0.002226	0.006742	0.318808
Robinson	0.020688	0.013678	0.008646	0.021083	0.137317
Albers Equal-Area Conic	0.021062	0.001036	0.004016	0.000374	0.025345
Mercator	0.022039	0.015430	0.003253	0.008891	0.598142
Lambert Conformal Conic	0.022039	0.001254	0.003253	0.000328	0.016815
Gall-Peters	0.022569	0.013472	0.020865	0.050588	0.491939
Polyconic	0.024400	0.001879	0.001503	0.000795	**0.012862**
Equirectangular	0.056594	0.014984	0.006930	0.006612	0.131018

Table 16.2: Regime-wise argmin extracted from Table 16.1.

Regime	Best projection	IDI (bits)
Global	Aitoff	0.015947
Regional mid-latitude	Lambert Azimuthal Equal-Area	0.000314
Regional equatorial	Aitoff	0.000445
Regional high-latitude	Lambert Azimuthal Equal-Area	0.000237
Arctic (75–90N)	Polyconic	0.012862

if anything, a strictly stronger illustration of Theorem 14.2's conclusion than the original four-regime, two-winner table: extending the pool by five further projections did not produce a new candidate uniform minimizer, and extending the regime collection by a fifth, geographically extreme regime (the Arctic, entirely absent from the original four regimes' latitude coverage) introduced a *third* distinct winner rather than resolving in favour of one of the first two. Bonne is worth flagging explicitly: at 0.014943 bits on the Arctic regime it is close to, but does not beat, Polyconic's 0.012862 bits, so it does not become a fourth winner, but its strong performance across the mid-latitude, equatorial, and high-latitude regimes (visible in Table 16.1) shows the extended pool is not merely padding the table with weak performers.

Key Result

No projection among the seventeen tested attains the regime-wise minimum at all five regimes; the uniform-minimizer test of Definition 13.3 fails for this larger pool and larger regime collection, exactly as it failed for the original twelve-projection, four-regime case of Chapter 14. This directly closes the first open item of Section 14.4 for this specific extension.

16.4 What remains open

This chapter checked one particular enlargement — five more projections, one more regime — not the fully universal, pool- and regime-free statement Section 14.4 already flagged as requiring machinery (a continuity or genericity argument over the relevant infinite-dimensional spaces) not developed anywhere in this monograph. Nothing here rules out some pool of eigh-

teen, or a differently constructed regime, eventually producing a uniform minimizer; the extension checked here only makes that outcome look, empirically, less rather than more likely, as the number of distinct regime-wise winners *increased* (from two to three) rather than collapsing to one when the pool and regime collection grew.

Part III

Fisher–Rao Geometry of Map Projections

Chapter 17

The Fisher–Rao Geometry of the Categorical Model

Part II treated the Informational Distortion Index purely as a number: a Jensen–Shannon divergence between two histograms, compared regime by regime and projection by projection. This chapter takes seriously the observation already flagged in Section 1.1, that any such comparison of probability distributions has a natural home in *information geometry*. Rather than working with an arbitrary parametric family, this chapter fixes the specific statistical model implicit in the companion volume's own construction — a K-bin histogram of pairwise distances, i.e., a point of the categorical (multinomial) probability simplex — and develops its Fisher–Rao geometry exactly, in closed form, rather than appealing to the general theory of statistical manifolds without specializing it. Everything proved here is purely analytic; no executed pipeline is required until Chapter 18 puts the geometry developed here to work on the companion volume's own histograms.

17.1 The categorical model and its interior

Definition 17.1 (The open categorical simplex)**.** For an integer $K \geq 2$, let

$$\Delta_K^\circ \;=\; \Big\{ p = (p_1, \ldots, p_K) \in \mathbb{R}^K \;:\; p_i > 0 \text{ for all } i,\; \textstyle\sum_{i=1}^K p_i = 1 \Big\} \tag{17.1}$$

denote the *open* (interior) probability simplex on K categories: the set of categorical distributions assigning strictly positive probability to every one of K categories. Throughout this chapter and Chapters 18 and 19, $K = 60$, matching the common histogram bin count already fixed in Section 15.1.

A raw 60-bin histogram of pairwise distances, normalized to sum to 1, is a point of the *closed* simplex $\overline{\Delta}_K$, not necessarily its interior: any bin with zero count lies on the boundary. The Fisher information metric constructed in Section 17.2 below is defined only on Δ_K° (its coefficients diverge as any $p_i \to 0$, Remark 17.3), so working with it requires moving every histogram strictly into the interior first.

Definition 17.2 (ε-regularization)**.** Given a raw histogram of non-negative integer bin counts $n = (n_1, \ldots, n_K)$ with $N = \sum_i n_i > 0$, and a fixed regularization constant $\varepsilon > 0$, the *ε-regularized histogram* is

$$p_i \;=\; \frac{n_i + \varepsilon}{N + K\varepsilon}, \qquad i = 1, \ldots, K. \tag{17.2}$$

Every p_i in Eq. (17.2) is strictly positive whenever $\varepsilon > 0$, regardless of how many bins n_i happen to be empty, and $\sum_i p_i = 1$ by construction; hence $p \in \Delta_K^\circ$ for every raw histogram and every $\varepsilon > 0$. As $\varepsilon \to 0^+$, $p_i \to n_i/N$, so Eq. (17.2) recovers the ordinary empirical frequency in the limit; the same construction, with $\varepsilon = 10^{-12}$, is already used (without this name) by the companion volume's own Jensen–Shannon divergence computation, which adds ε to each

bin before normalizing for exactly the same reason — to avoid a $\log 0$ term in the Kullback–Leibler divergence. This chapter reuses the identical $\varepsilon = 10^{-12}$ throughout, so that the JSD recomputed here from ε-regularized histograms agrees with the companion pipeline's own JSD values to the precision already reported in Table 15.1, rather than introducing a second, silently different convention.

Remark 17.3 (Why the boundary is excluded, not just avoided). The exclusion of the boundary $\partial\overline{\Delta}_K = \overline{\Delta}_K \setminus \Delta_K^\circ$ is not a technical nicety adopted only for this monograph's convenience: the Fisher information metric of Definition 17.4 below has a coordinate coefficient $1/p_i$ that diverges as $p_i \to 0^+$, so Δ_K° genuinely is the natural domain of the geometric structure developed in this chapter, and no boundary extension of that structure is attempted here (Section 17.8 returns to this point explicitly).

17.2 The Fisher information metric of the categorical model

Definition 17.4 (Fisher information metric, categorical model)**.** Parametrize Δ_K° by its own points $p = (p_1, \dots, p_K)$ (the categorical model is the trivial case of a statistical model in which the distribution *is* the parameter). The Fisher information metric g_{Fisher} on Δ_K° is the Riemannian metric whose value at p, on tangent vectors δ, δ' tangent to the simplex (i.e., $\sum_i \delta_i = \sum_i \delta'_i = 0$), is

$$g_{\text{Fisher}}(\delta, \delta')\big|_p \;=\; \sum_{i=1}^{K} \frac{\delta_i\, \delta'_i}{p_i}. \tag{17.3}$$

Equation (17.3) is the classical Fisher information metric of the multinomial family, specialized to the case in which the model parameters are the cell probabilities themselves (Rao, 1945): for a general parametric family p_θ, the Fisher information matrix is $I_{jk}(\theta) = \mathbb{E}_{p_\theta}\big[\partial_j \log p_\theta \cdot \partial_k \log p_\theta\big]$, and for the categorical model with $\theta = p$ this reduces, by direct computation ($\partial_i \log p_i = 1/p_i$ on the i-th coordinate density, zero on the others, so $I_{ii}(p) = \mathbb{E}_p[(1/p_i)^2 \mathbf{1}_{\{i\}}] = p_i \cdot (1/p_i)^2 = 1/p_i$ and $I_{ij}(p) = 0$ for $i \neq j$), to exactly the diagonal quadratic form of Eq. (17.3). This is a standard fact of information geometry (see also Kass, 1989, and Amari & Nagaoka, 2000, for the general theory this specializes); it is derived here from first principles, not merely cited, because the rest of this chapter depends on its exact coefficients.

17.3 The square-root map and its isometry to a sphere

Definition 17.5 (The square-root map)**.** Let $\varphi : \Delta_K^\circ \to \mathbb{R}^K$ be defined coordinatewise by

$$\varphi(p) \;=\; \big(2\sqrt{p_1}, \dots, 2\sqrt{p_K}\big). \tag{17.4}$$

Since $\sum_i p_i = 1$, every image point satisfies $\sum_i \varphi(p)_i^2 = 4\sum_i p_i = 4$, so φ maps Δ_K° into the open positive orthant of the sphere $S^{K-1}(2) = \{u \in \mathbb{R}^K : |u| = 2\}$ of radius 2 centred at the origin; write $S_+^{K-1}(2)$ for this open orthant-intersected sphere. φ is a smooth bijection onto its image, with smooth inverse $\varphi^{-1}(u) = (u_1^2/4, \dots, u_K^2/4)$.

Theorem 17.6 (The square-root map is an isometry onto the round sphere)**.** $\varphi : \big(\Delta_K^\circ, g_{\text{Fisher}}\big) \to \big(S_+^{K-1}(2), g_{\text{round}}\big)$ *is a Riemannian isometry, where* g_{round} *is the metric induced on the sphere by the ambient Euclidean metric of* $\mathbb{R}^K$.

Proof. Let δ be a tangent vector to Δ_K° at p (so $\sum_i \delta_i = 0$). Differentiating $\varphi_i(p) = 2\sqrt{p_i}$ coordinatewise gives $d\varphi_i(\delta) = \delta_i/\sqrt{p_i}$. The ambient Euclidean (hence round, since the sphere carries the metric induced from the ambient space) inner product of the pushed-forward vectors $d\varphi(\delta), d\varphi(\delta')$ is

$$\langle d\varphi(\delta), d\varphi(\delta')\rangle \;=\; \sum_{i=1}^{K} \frac{\delta_i}{\sqrt{p_i}} \cdot \frac{\delta_i'}{\sqrt{p_i}} \;=\; \sum_{i=1}^{K} \frac{\delta_i \delta_i'}{p_i} \;=\; g_{\text{Fisher}}(\delta, \delta')\big|_p,$$

by Eq. (17.3). Since $d\varphi$ is a linear isomorphism onto the tangent space of the sphere at $\varphi(p)$ (the tangent space to $S^{K-1}(2)$ at u is exactly $\{v : \sum_i u_i v_i = 0\}$, and $\sum_i \varphi_i(p)\, d\varphi_i(\delta) = \sum_i 2\sqrt{p_i}\cdot \delta_i/\sqrt{p_i} = 2\sum_i \delta_i = 0$ confirms $d\varphi(\delta)$ indeed lies in this tangent space), $d\varphi$ is a linear isometry of tangent spaces at every point, which is precisely the definition of a Riemannian isometry. □

Remark 17.7 (This is Rao's original construction, made explicit). The isometry of Theorem 17.6 is the original geometric content of Rao (1945)'s introduction of the Fisher information metric for the multinomial family, stated there in different notation; it is re-derived here in full, coordinate-by-coordinate, because Chapters 18 and 19 use its exact constants (the radius 2, not 1 or some other normalization) directly, and a citation without the explicit computation would leave those later chapters' formulas unmotivated.

17.4 The Fisher–Rao distance in closed form

Corollary 17.8 (Fisher–Rao distance, Bhattacharyya coefficient, Hellinger distance)**.** *For $p, q \in \Delta_K^\circ$, let*

$$\mathrm{BC}(p, q) \;=\; \sum_{i=1}^{K} \sqrt{p_i q_i} \tag{17.5}$$

denote the Bhattacharyya coefficient *of p and q (so $\mathrm{BC}(p,q) \in (0,1]$, with equality to 1 iff $p = q$, by the Cauchy–Schwarz inequality applied to $(\sqrt{p_i})_i, (\sqrt{q_i})_i$, both unit vectors in the ℓ^2 norm). Then the Fisher–Rao geodesic distance between p and q, i.e., the length of the shortest path in $(\Delta_K^\circ, g_{\text{Fisher}})$ joining them, is*

$$d_{\text{FR}}(p, q) \;=\; 2 \arccos\big(\mathrm{BC}(p, q)\big), \tag{17.6}$$

and the (squared) Hellinger distance $H(p,q)^2 = \frac{1}{2}\sum_i (\sqrt{p_i} - \sqrt{q_i})^2$ satisfies $H(p,q) = \sqrt{1 - \mathrm{BC}(p,q)}$.

Proof. By Theorem 17.6, $d_{\text{FR}}(p,q)$ equals the round-metric geodesic distance between $u = \varphi(p)$ and $v = \varphi(q)$ on the sphere $S^{K-1}(2)$, which is the standard great-circle formula $d_{\text{FR}}(p,q) = 2\arccos(u \cdot v/4)$ (radius-2 sphere; the great-circle distance between two points of a sphere of radius r separated by angle ω is $r\omega$, and $\omega = \arccos(u \cdot v / r^2)$). Since $u \cdot v = \sum_i \varphi_i(p)\varphi_i(q) = \sum_i 4\sqrt{p_i q_i} = 4\,\mathrm{BC}(p,q)$, this is exactly Eq. (17.6). For the Hellinger identity, expand directly:

$$H(p,q)^2 = \tfrac{1}{2}\sum_i \big(p_i + q_i - 2\sqrt{p_i q_i}\big) = \tfrac{1}{2}\big(1 + 1 - 2\,\mathrm{BC}(p,q)\big) = 1 - \mathrm{BC}(p,q). \qquad \square$$

This is exactly the formula already anticipated in Chapter 18's construction, with $\omega = \arccos(\mathrm{BC}(p,q))$ there matching $\arccos(u \cdot v/4)$ here, and it is what `code/fisher_rao_pipeline.py` computes directly (Section 19.1).

17.5 Constant curvature

Corollary 17.9 (Constant positive curvature)**.** $(\Delta_K^\circ, g_{\text{Fisher}})$ *has constant Gaussian* (*sectional*) *curvature* $1/4$*.*

Proof. By Theorem 17.6, $(\Delta_K^\circ, g_{\text{Fisher}})$ is isometric to (an open subset of) the round sphere $S^{K-1}(2)$ of radius 2 in Euclidean space, and a round sphere of radius r has constant sectional curvature $1/r^2$ (a standard fact of Riemannian geometry). With $r = 2$, this is $1/4$. □

17.6 Local relation to the Jensen–Shannon divergence

The companion volume's Informational Distortion Index is a Jensen–Shannon divergence, not a Fisher–Rao distance; this section makes explicit exactly how the two are related when p and q are close, and exactly why closeness is needed for that relation to hold.

Theorem 17.10 (Local quadratic approximation of the JSD by the Fisher–Rao distance)**.** *Let* $p \in \Delta_K^\circ$ *and let* $q = p + \delta$ *with* $\sum_i \delta_i = 0$ *and* $\delta \to 0$*. Then, to leading* (*quadratic*) *order in* δ*,*

$$\text{JSD}_{\text{nats}}(p, q) \;=\; \frac{1}{8}\, d_{\text{FR}}(p, q)^2 \;+\; O(|\delta|^3), \tag{17.7}$$

where $\text{JSD}_{\text{nats}} = (\ln 2)\,\text{JSD}_{\text{bits}}$ *is the Jensen–Shannon divergence measured in nats* (*natural logarithm*) *rather than bits.*

Proof. Write $m = (p+q)/2 = p+\delta/2$. A second-order Taylor expansion of the Kullback–Leibler divergence around coincident arguments gives, for any a close to b (both in Δ_K°), $D_{\text{KL}}(a\|b) = \frac{1}{2}\sum_i (a_i - b_i)^2/b_i + O(|a-b|^3)$ (the first-order term vanishes since $D_{\text{KL}}(a\|a) = 0$ is a minimum). Apply this with $(a, b) = (p, m)$ and $(a, b) = (q, m)$: since $p - m = \delta/2$ and $q - m = -\delta/2$, and $m = p + O(\delta)$ so $1/m_i = 1/p_i + O(\delta)$,

$$D_{\text{KL}}(p\|m) = \tfrac{1}{2}\sum_i \frac{(\delta_i/2)^2}{p_i} + O(|\delta|^3) = \tfrac{1}{8}\sum_i \frac{\delta_i^2}{p_i} + O(|\delta|^3),$$

and identically $D_{\text{KL}}(q\|m) = \frac{1}{8}\sum_i \delta_i^2/p_i + O(|\delta|^3)$ (the two agree to this order by the symmetry $p - m = -(q - m)$ up to the $O(\delta)$ correction to $1/m_i$, which only affects the $O(|\delta|^3)$ remainder). Hence

$$\text{JSD}_{\text{nats}}(p, q) = \tfrac{1}{2}\big[D_{\text{KL}}(p\|m) + D_{\text{KL}}(q\|m)\big] = \tfrac{1}{8}\sum_i \frac{\delta_i^2}{p_i} + O(|\delta|^3) = \tfrac{1}{8}\, g_{\text{Fisher}}(\delta, \delta)\big|_p + O(|\delta|^3),$$

by Eq. (17.3). On the other hand, for q in a small geodesic neighbourhood of p, the Riemannian distance satisfies $d_{\text{FR}}(p, q)^2 = g_{\text{Fisher}}(\delta, \delta)\big|_p + O(|\delta|^3)$ exactly (a general fact for any Riemannian metric: geodesic distance squared agrees with the metric's own quadratic form on the tangent space to leading order, since the metric coefficients are themselves smooth and $q - p = \delta$ to leading order along the geodesic). Combining the two displays gives Eq. (17.7). □

Remark 17.11 (What Theorem 17.10 does, and does not, say about far-apart histograms). Theorem 17.10 is a strictly local (infinitesimal-neighbourhood) statement: it says nothing, by itself, about the relation between $\text{JSD}_{\text{nats}}(p, q)$ and $d_{\text{FR}}(p, q)^2/8$ when p, q are not close — and the companion volume's own histograms, compared at the Global regime, are generally not close in this sense (Table 18.1's Fisher–Rao distances range up to 0.6 radians, far from infinitesimal). Section 19.5 and Section 17.7 below address the far-apart case directly, and separately from this local result.

17.7 A global sandwich bound between the Fisher–Rao and Hellinger quantities

Beyond the local relation of Theorem 17.10, it is natural to ask whether $\mathrm{JSD}_{\mathrm{nats}}$ and the Hellinger distance H of Corollary 17.8 satisfy some inequality valid for *all* $p, q \in \Delta_K^\circ$, not just infinitesimally close pairs. The following bound is stated as a proposition rather than a theorem, deliberately: it is reported here exactly as it was found, verified numerically rather than proved analytically, consistent with this monograph's discipline of not blurring the line between the two.

Proposition 17.12 (A global JSD–Hellinger sandwich, numerically verified)**.** *For all* $p, q \in \Delta_K^\circ$,

$$\tfrac{1}{2} H(p,q)^2 \;\le\; \mathrm{JSD}_{\mathrm{nats}}(p,q) \;\le\; H(p,q). \tag{17.8}$$

Provenance: a synthetic, executed numerical check, not a proof

Proposition 17.12 is not proved in this monograph. It was checked numerically on 20,000 independent pairs (p, q) drawn as $p, q \sim \mathrm{Dirichlet}(\alpha \mathbf{1}_{60})$ for α drawn uniformly from $[0.05, 5]$ independently for each pair (a range chosen to span both highly concentrated and nearly-uniform categorical distributions over $K = 60$ categories), each pair ε-regularized exactly as in Definition 17.2, using the fixed random seed 42; the bound of Eq. (17.8) held, without a single exception (to a numerical tolerance of 10^{-9}, accounting for floating-point round-off), on all 20,000 synthetic pairs. Section 19.5 of Chapter 19 reruns this same check on 48 genuine cartographic pairs — the companion pipeline's own projection histograms against their regime's geodesic reference — as an independent, structurally different test; it also holds there without exception. Neither check is a proof: both are consistent with Eq. (17.8) holding universally on Δ_K°, and consistent with it merely holding on every pair tested so far, and this monograph does not claim to know which of the two is the case. A short heuristic can be offered for why the bound is at least plausible in the direction it is verified: both sides of Eq. (17.8) vanish together (at $p = q$) and, by Theorem 17.10, agree with each other to within a factor of 2 at leading order near $p = q$ (since $H(p,q)^2 \approx \frac{1}{4} d_{\mathrm{FR}}(p,q)^2$ for close p, q, itself a short computation from Corollary 17.8's $H^2 = 1 - \mathrm{BC}$ and $\mathrm{BC} = \cos(d_{\mathrm{FR}}/2) \approx 1 - d_{\mathrm{FR}}^2/8$), which is consistent with, though does not establish, an inequality holding globally.

17.8 What this chapter establishes, and what remains open

This chapter has fixed the specific statistical model left implicit in the companion volume's own construction — the $K = 60$-bin categorical histogram, restricted to the open simplex Δ_K° by the ε-regularization of Definition 17.2 — and developed its Fisher–Rao geometry in closed form: the square-root map is an exact isometry onto a round sphere of radius 2 (Theorem 17.6), giving a closed-form Fisher–Rao distance in terms of the Bhattacharyya coefficient (Corollary 17.8) and constant curvature $1/4$ (Corollary 17.9); the Jensen–Shannon divergence used throughout Part II is exactly the leading-order quadratic approximation to (one-eighth of) the squared Fisher–Rao distance in the local, infinitesimal-neighbourhood regime (Theorem 17.10), a relation that does *not* extend, as a matter of proof, to the far-apart regime in which the companion volume's own histograms actually sit. Left explicitly open: a global, analytically proved version of the sandwich bound of Proposition 17.12, verified here only numerically; the behaviour of the geometry at the excluded boundary $\partial\overline{\Delta}_K$ of the simplex, where

the Fisher metric's coefficients diverge (Remark 17.3) and where a zero-count histogram bin genuinely lands before ε-regularization moves it into the interior; and any treatment of statistical models on the simplex other than the categorical model considered here (a different choice of model — e.g., treating the 60 bin *boundaries* rather than the bin *probabilities* as the parameters — would induce a different Fisher metric, not considered in this monograph). Chapter 18 now puts this geometry to work directly on the companion volume's own twelve-projection pool.

Chapter 18

Map Projections as Points on a Statistical Manifold

Chapter 17 developed the Fisher–Rao geometry of the categorical model in the abstract, for an arbitrary pair of histograms $p, q \in \Delta_K^\circ$. This chapter puts that geometry to work on the specific histograms already computed in Chapter 15: for each of the four sampling regimes of Section 15.1 and each of the twelve projections of Theorem 14.2, the (median-rescaled, ε-regularized) projected-distance histogram is a single point of Δ_{60}°, and the geodesic-distance histogram is a thirteenth, distinguished point — the one every projection's histogram is compared against. Reading the whole projection pool this way turns Table 15.1's table of numbers into an explicit configuration of points on a curved statistical manifold, with its own distances, geodesics, and structure beyond what the JSD column-by-column comparison shows.

18.1 The embedding of a projection pool into the manifold

Definition 18.1 (The manifold point of a projection under a regime)**.** Fix a sampling regime $R = (\Omega, G) \in \mathcal{R}$ in the sense of Definition 13.1, with $|G| = n$ points and $\binom{n}{2}$ pairwise distances binned into the $K = 60$ common histogram bins of Section 15.1. For a projection $f \in \mathcal{P}$, let $\pi_R(f) \in \Delta_{60}^\circ$ denote the ε-regularized, normalized projected-distance histogram of f under R (Definition 17.2), and let $\gamma_R \in \Delta_{60}^\circ$ denote the corresponding ε-regularized geodesic-distance histogram — the same object underlying every JSD entry of Table 15.1's row for R. Definition 13.2's IDI is then exactly

$$\mathrm{IDI}(f, R) \;=\; \mathrm{JSD}_{\mathrm{bits}}\big(\gamma_R,\ \pi_R(f)\big), \tag{18.1}$$

and this chapter's Fisher–Rao analogue is $d_{\mathrm{FR}}(\gamma_R, \pi_R(f))$, computed via Eq. (17.6) from bit-for-bit the same two histograms.

For a fixed regime R, the map $f \mapsto \pi_R(f)$ sends the twelve-projection pool $\mathcal{P}$, together with the distinguished reference point γ_R, to thirteen points of the sphere $S_+^{59}(2)$ (via Theorem 17.6's isometry φ) — a configuration on which every notion of Chapter 17 (distance, geodesic, curvature) applies directly.

Provenance of the histograms used in this chapter

Every histogram $\pi_R(f)$ and γ_R used below is read directly from the same executed pipeline as Chapter 15: `code/fisher_rao_pipeline.py` imports the companion volume's own grid-construction, haversine-distance, and twelve-projection routines from `informational_distortion.py` without modification, and adds only the Bhattacharyya-coefficient and Fisher–Rao-distance computations of Corollary 17.8. No histogram is recomputed by a different method than the one already used for Table 15.1; the JSD values recovered by this chapter's own script agree with Table 15.1 to the reported precision, which is itself a direct executed cross-check that the two pipelines' shared components

(grids, projections, histograms) are bit-for-bit identical, not merely similar.

18.2 The pairwise Fisher–Rao distance matrix, Global regime

Table 18.1 reports the full 12×12 pairwise Fisher–Rao distance matrix among the twelve projections' own Global-regime histograms $\pi_{\text{Global}}(f)$ (not against γ_{Global}, which is the subject of Section 19.2 of Chapter 19 instead) — the first genuinely new object this chapter reports, since none of Tables 15.1 and 15.2's entries compare two projections directly to each other.

Table 18.1: Pairwise Fisher–Rao distance d_{FR} (radians) between the twelve projections' own Global-regime histograms, direct output of `code/fisher_rao_pipeline.py`. The exact zero entry (Mer–LCC) marks projections whose 60-bin histograms are bit-for-bit identical (Proposition 18.2).

	Ait	Ham	Mol	Win	Sin	LAEA	VdG	Rob	Alb	Mer	LCC	Equ
Aitoff (Ait)	–	0.053	0.080	0.116	0.153	0.118	0.135	0.125	0.204	0.226	0.226	0.522
Hammer (Ham)	0.053	–	0.082	0.128	0.145	0.123	0.151	0.129	0.213	0.238	0.238	0.531
Mollweide (Mol)	0.080	0.082	–	0.112	0.186	0.132	0.134	0.136	0.200	0.227	0.227	0.510
Winkel Tripel (Win)	0.116	0.128	0.112	–	0.169	0.140	0.130	0.124	0.199	0.221	0.221	0.527
Sinusoidal (Sin)	0.153	0.145	0.186	0.169	–	0.175	0.213	0.152	0.245	0.292	0.292	0.574
Lambert Azim. E.-A. (LAEA)	0.118	0.123	0.132	0.140	0.175	–	0.159	0.142	0.224	0.237	0.237	0.535
Van der Grinten (VdG)	0.135	0.151	0.134	0.130	0.213	0.159	–	0.115	0.216	0.205	0.205	0.486
Robinson (Rob)	0.125	0.129	0.136	0.124	0.152	0.142	0.115	–	0.211	0.216	0.216	0.513
Albers Equal-Area Conic (Alb)	0.204	0.213	0.200	0.199	0.245	0.224	0.216	0.211	–	0.286	0.286	0.533
Mercator (Mer)	0.226	0.238	0.227	0.221	0.292	0.237	0.205	0.216	0.286	–	0.000	0.525
Lambert Conformal Conic (LCC)	0.226	0.238	0.227	0.221	0.292	0.237	0.205	0.216	0.286	0.000	–	0.525
Equirectangular (Equ)	0.522	0.531	0.510	0.527	0.574	0.535	0.486	0.513	0.533	0.525	0.525	–

Table 18.1 shows a clear structural pattern, echoing (but not identical to) the equal-area/conformal/compromise groupings already familiar from the companion volume's own qualitative classification of these projections: Aitoff, Hammer, and Mollweide form the single tightest triple in the whole matrix, with mutual distances between 0.053 and 0.082; Winkel Tripel and Robinson sit at moderate distance (0.11–0.14) from this triple and from each other; and Mercator and Lambert Conformal Conic form an exactly coincident pair (Proposition 18.2) that nonetheless sits at moderate-to-large distance from every other projection (0.205 to 0.292). Equirectangular stands apart from every other projection in the pool: every one of its eleven distances lies in $[0.486, 0.574]$, while every distance between two *non*-Equirectangular projections lies in $[0, 0.292]$ — Equirectangular's closest neighbour in the whole pool (Van der Grinten, at 0.486) is still farther from it than the two most mutually distant non-Equirectangular projections (Sinusoidal and Mercator/LCC, at 0.292) are from each other, a genuine, unambiguous separation rather than a matter of degree. This is exactly the outlier role Equirectangular already played in Table 15.1's Global column, where its IDI value (0.056594 bits) is more than double the next-worst projection's.

18.3 An exact coincidence: Mercator and Lambert Conformal Conic

Proposition 18.2 (Mercator and Lambert Conformal Conic coincide exactly, at two regimes)**.** *At the Global and Regional equatorial regimes, the* 60*-bin histograms* $\pi_R(\text{Mercator})$ *and* $\pi_R(\text{LCC})$ *are identical bin-for-bin (hence* $d_{\text{FR}} = 0$ *and* $\text{JSD} = 0$ *exactly between them); at the Regional mid-latitude and Regional high-latitude regimes they are not* $(d_{\text{FR}}(\pi_R(\text{Mercator}), \pi_R(\text{LCC})) = 0.323312$ *and* 0.233718 *respectively, computed directly by the same script).*

This sharpens, into an exact statement about the underlying histograms rather than only their JSD summary, the observation already flagged honestly in Section 15.3: that Mercator and

Lambert Conformal Conic report identical JSD values at the Global and Regional equatorial regimes is not a coincidence of six-decimal rounding, but the two projections' pairwise-distance histograms being *literally the same finite object*, verified here directly by the pipeline's own zero-distance output (Table 18.1's Mer–LCC entry) rather than inferred indirectly from equal JSD values. The mechanism is the parametrization convention already documented in Section 15.1: Lambert Conformal Conic is fitted with standard parallels at the sample's own $25^{\text{th}}/75^{\text{th}}$ latitude percentiles, and at the Global and Regional-equatorial grids specifically, this fitted conic construction reduces, for the particular percentile values these two grids happen to produce, to a projection whose pairwise planar distances coincide exactly with Mercator's own — while at the Regional mid-latitude and high-latitude grids, whose latitude percentiles differ, the coincidence does not persist. No claim is made here about why the coincidence holds at exactly these two regimes and not the other two beyond this parametrization-level observation; a full explanation would require tracing the conic projection's scale-factor formula symbolically as a function of the fitted standard parallels, which is not attempted.

18.4 Explicit geodesics on the statistical manifold

Theorem 17.6 makes the geodesics of $(\Delta_K^\circ, g_{\text{Fisher}})$ completely explicit: they are the images, under φ^{-1}, of great-circle arcs on the sphere $S_+^{K-1}(2)$.

Definition 18.3 (Fisher–Rao geodesic interpolation)**.** For $p, q \in \Delta_K^\circ$ with $u = \varphi(p)$, $v = \varphi(q)$, and angular separation $\omega = \arccos(u \cdot v/4)$ (so $d_{\text{FR}}(p, q) = 2\omega$), the Fisher–Rao geodesic from p to q is $t \mapsto \varphi^{-1}(w(t))$ for $t \in [0, 1]$, where

$$w(t) \;=\; \frac{\sin\big((1-t)\omega\big)}{\sin\omega}\,u \;+\; \frac{\sin(t\omega)}{\sin\omega}\,v \tag{18.2}$$

is the standard spherical-interpolation ("slerp") formula for the great-circle arc between two points of a sphere.

Equation (18.2) is the unique constant-speed parametrization of the great circle through u, v on the sphere of radius 2 (a standard fact of spherical geometry: $w(t)$ lies on the sphere since $|w(t)| = 2$ for all t by direct computation using $u \cdot u = v \cdot v = 4$, and traces the shorter great-circle arc between u and v at constant angular speed ω); by Theorem 17.6's isometry, $t \mapsto \varphi^{-1}(w(t))$ is therefore the Fisher–Rao geodesic in Δ_K°, and in particular the midpoint $\varphi^{-1}(w(1/2))$ is the Fisher–Rao *midpoint* (Fréchet mean along the geodesic) of p and q.

Executed additivity check

Definition 18.3's construction was verified directly, not merely asserted, on two pairs drawn from the Global regime's own projection pool: (Aitoff, Lambert Azimuthal Equal-Area) — the pair already singled out in Theorem 14.2's proof as instantiating Proposition 14.1's adversarial hypothesis — and (Mercator, Lambert Conformal Conic) — the exactly-coincident pair of Proposition 18.2. For each pair, the midpoint $w(1/2)$ was computed via Eq. (18.2), mapped back to a histogram via φ^{-1}, and the Fisher–Rao distances from each endpoint to this midpoint were summed and compared against the direct endpoint-to-endpoint distance $d_{\text{FR}}(p, q)$. For the Aitoff/LAEA pair, $d_{\text{FR}}(p, q) = 0.11784487$ against a sum-of-halves of 0.11784487, an additivity gap of 7.9×10^{-12} (floating-point round-off, not a modelling discrepancy); for the exactly-coincident Mercator/LCC pair, both quantities are 0 exactly, as they must be for two identical points. This confirms that Eq. (18.2)'s construction genuinely traces the metric's own geodesic (additivity of distance

along a geodesic is a defining property a merely illustrative interpolation would have no reason to satisfy to 10^{-12} precision), not an arbitrary path between the two histograms that happens to look reasonable.

18.5 What a geodesic between two projections means, and does not mean

An honest reading of Definition 18.3

The Fisher–Rao geodesic between π_R(Aitoff) and π_R(Lambert Azimuthal Equal-Area) is a path through the space of *distance-histograms*, each point of which is, by Eq. (18.2), a genuine 60-bin categorical distribution — but it is emphatically **not** claimed here that every point along this path corresponds to the pairwise-distance histogram of some actual, interpolated map projection "between" Aitoff and Lambert Azimuthal Equal-Area in any cartographic sense. The statistical manifold of Chapter 17 is a geometric structure on the *space of histograms that projections happen to produce*, not on the space of projections itself (which has its own, entirely different structure — a projection is a smooth map $S^2 \to \mathbb{R}^2$, not a probability distribution); two different projections could, in principle, produce numerically close or identical histograms (as Mercator and Lambert Conformal Conic in fact do, Proposition 18.2) while being, as maps, genuinely different constructions. The geodesic of Definition 18.3 is therefore a rigorous statement about interpolating between two *distributions*, useful for the numerical comparisons of Chapter 19, and no stronger geometric or cartographic claim about the projections themselves is made or implied.

18.6 What this chapter proves, and what remains open

This chapter has computed, directly from the executed pipeline, the full pairwise Fisher–Rao distance structure of the companion volume's twelve-projection pool at the Global regime (Table 18.1); proved an exact, non-asymptotic coincidence between two specific projections at two specific regimes (Proposition 18.2), traceable to their shared fitted-parallel construction; given an explicit, verified construction of Fisher–Rao geodesics between any two histograms (Definition 18.3, checked by the additivity property to 10^{-12} precision); and stated plainly the limits of what such a geodesic does and does not represent (Section 18.5). Left open: a systematic account of the pairwise-distance-matrix structure at the three regional regimes (only the Global-regime matrix is reported in full, Table 18.1, though the per-regime pairwise Fisher–Rao data underlying such an account is already present in `data/fisher_rao_results.json`); any formal clustering analysis (hierarchical or otherwise) of the pairwise distance matrix, in the style of the hierarchical-clustering analysis of the companion volume's four-axis chapter (Chapter 15), which is not attempted here; and any characterization of which projections, as a general rule rather than the one specific instance of Proposition 18.2, produce coincident or near-coincident histograms. Chapter 19 now turns from the pool's internal structure to its relationship with the geodesic reference point γ_R across all four regimes, closing the loop back to Theorem 14.2's impossibility theorem.

Chapter 19

Numerical Illustration: Fisher–Rao Distances Across the Companion Volume's Pool

This chapter reports the direct, executed computation underlying Chapter 18's discussion, extended now to all four sampling regimes of Section 15.1 rather than only the Global regime of Section 18.2. Its central question is a genuinely new one, not addressed anywhere in Chapters 13 to 15: does Theorem 14.2's impossibility conclusion, and in particular its specific regime-wise minimizers, survive being re-measured with the Fisher–Rao distance of Corollary 17.8 in place of the Jensen–Shannon divergence used throughout Part II?

19.1 Methodology and setup

Every number in this chapter is computed by `code/fisher_rao_pipeline.py`, which imports the companion volume's own grid-construction, haversine-distance, and twelve-projection routines from `informational_distortion.py` without modification (the same import already documented for Chapter 15's own pipeline), and computes, for every (projection, regime) pair, the Bhattacharyya coefficient BC, the Hellinger distance H, and the Fisher–Rao distance d_{FR} (Eq. (17.6)) between that projection's histogram and the regime's geodesic-distance histogram, alongside the JSD already reported in Table 15.1. All four regimes — Global, Regional mid-latitude, Regional equatorial, Regional high-latitude — are exactly those of Section 15.1, re-executed rather than re-typed from that chapter's table.

What is, and is not, re-derived in this chapter

The JSD column reported alongside d_{FR} below is recomputed by this chapter's own script, not copied from Table 15.1; the two agree to the reported precision at every entry (a direct executed cross-check that the shared grid, projection, and histogram code invoked by both scripts is bit-for-bit identical), and this agreement is the reason Section 19.3 below can treat the JSD ranking of Table 15.2 and this chapter's Fisher–Rao ranking as computed from literally the same underlying histograms, differing only in which divergence is applied to them.

19.2 Full results

Table 19.1: Fisher–Rao distance d_{FR} (radians, Eq. (17.6)) between each projection's histogram and the geodesic reference histogram, all twelve projections, all four regimes. Boldface marks the regime-wise minimum. Compare directly against Table 15.1's JSD table, computed from the same underlying histograms.

Projection	Global	Mid-lat.	Equatorial	High-lat.
Aitoff	**0.299672**	0.297668	**0.049686**	0.237727
Hammer	0.301667	0.256846	0.055834	0.264175
Mollweide	0.306632	0.206564	0.184228	0.263813
Winkel Tripel	0.312717	0.231408	0.192075	0.274570
Sinusoidal	0.315205	0.126293	0.072772	0.185869
Lambert Azimuthal Equal-Area	0.331553	**0.041736**	0.073120	**0.036239**
Van der Grinten	0.333664	0.253766	0.111141	0.198002
Robinson	0.343627	0.282466	0.219646	0.348748
Albers Equal-Area Conic	0.346679	0.075820	0.149369	0.045562
Mercator	0.356702	0.299415	0.134404	0.232635
Lambert Conformal Conic	0.356702	0.083408	0.134404	0.042800
Equirectangular	0.605068	0.295659	0.196294	0.202165

19.3 Cross-validating the impossibility theorem under a different metric

Proposition 19.1 (The regime-wise minimizer is metric-independent, for this pool)**.** *At every one of the four regimes of Theorem 14.2, the projection minimizing d_{FR} (Table 19.1) is the same projection that minimizes the JSD (Table 15.1): Aitoff at Global and Regional equatorial, Lambert Azimuthal Equal-Area at Regional mid-latitude and Regional high-latitude.*

This is read off Table 19.1 directly (boldface entries), and is not a logical consequence of Theorem 17.10's local approximation — the histograms compared here are not infinitesimally close, so Eq. (17.7) does not, by itself, force the two divergences' minimizers to coincide. Proposition 19.1 is reported here as a further executed check, not a corollary of anything already proved: it confirms that Theorem 14.2's specific adversarial structure (Aitoff and Lambert Azimuthal Equal-Area trading off across regimes) is not an artefact of the particular choice of Jensen–Shannon divergence as the comparison functional, but persists when the comparison is redone with a genuinely different, Riemannian, divergence built from the same histograms.

What Proposition 19.1 does not show

Proposition 19.1 is a statement about *regime-wise minimizers* only — the single entry Theorem 14.2's proof actually needs. It is emphatically not a statement that the two divergences induce the *same full ranking* of all twelve projections at every regime: Section 19.4 shows directly that they do not.

19.4 Where the two full rankings disagree

Ordering all twelve projections by d_{FR} at each regime and comparing against the corresponding ordering by JSD (Table 15.1) gives identical orderings at the Global and Regional equatorial regimes, and exactly one adjacent transposition at each of the other two, reported in full in Table 19.2.

Table 19.2: The only two positions, across all four regimes, at which the Fisher–Rao ranking and the JSD ranking of the twelve-projection pool disagree. All other ten ranks at each of these two regimes, and all twelve ranks at the Global and Regional equatorial regimes, agree exactly (not shown).

Regime	Rank	By d_{FR}	By JSD
Regional mid-latitude	11th	Aitoff (0.297668)	Mercator (0.015430 bits)
Regional mid-latitude	12th	Mercator (0.299415)	Aitoff (0.015887 bits)
Regional high-latitude	5th	Van der Grinten (0.198002)	Equirectangular (0.006612 bits)
Regional high-latitude	6th	Equirectangular (0.202165)	Van der Grinten (0.006742 bits)

Both disagreements are transpositions of a near-tie: at Regional mid-latitude, Aitoff's JSD (0.015887 bits) exceeds Mercator's (0.015430 bits) by only 0.000457 bits, a difference of about 3% relative to either value, while their Fisher–Rao distances (0.297668 against 0.299415) place them in the opposite order by an even smaller relative margin; at Regional high-latitude, Van der Grinten and Equirectangular are separated by only 0.000130 bits in JSD but swap order under d_{FR}. Both instances occur among the *worst*-performing projections at their respective regimes (ranks 11–12 and 5–6 of 12, the latter middling rather than extreme), never among the best few, and in particular never disturb Proposition 19.1's regime-wise minimizer. The honest reading is that Fisher–Rao distance and Jensen–Shannon divergence, while both derived from and vanishing together on the same pair of histograms (Section 17.6), are not monotonic transformations of one another in general — Proposition 17.12's sandwich bounds d_{FR} between two functions of JSD, not to a single function of it, so occasional rank swaps among closely tied projections are exactly what the theory of Chapter 17 predicts, not an inconsistency to be explained away.

19.5 Confirming the global sandwich on genuine cartographic data

Proposition 17.12 of Chapter 17 was reported as numerically verified only on synthetic random distributions. Table 19.1's underlying 48 (projection, regime) pairs — twelve projections at each of four regimes, each compared against that regime's geodesic histogram — give a second, independent, and substantively different test: genuine cartographic histograms, not random simplex draws. The check $\frac{1}{2}H^2 \leq \mathrm{JSD}_{\mathrm{nats}} \leq H$ was re-run directly on all 48 real pairs (the same computation already flagged in Definition 17.2's data-provenance discussion); it holds, without exception, at every one of the 48 entries. This is worth stating plainly rather than assuming it would automatically follow from the synthetic-data check of Chapter 17: real cartographic distance histograms, being derived from actual projection formulas rather than Dirichlet draws, could in principle have concentrated much of their mass in a small number of bins in a way that stresses the sandwich bound differently from the broad synthetic distributions tested in Proposition 17.12's original verification — and indeed several of the regional-regime histograms are visibly concentrated, spanning a much narrower range of physical distances than the Global regime's (Section 15.3's first bulleted observation). That the bound still holds on this structurally different data is an additional, independent piece of numerical support for Proposition 17.12, not a restatement of the same check.

19.6 Closing the roadmap for Part III

This completes the three-chapter arc for Part III promised in Sections 1.1 and 1.2: Chapter 17 developed the Fisher–Rao geometry of the categorical model exactly, fixing the specific statistical model (the 60-bin categorical histogram) the companion volume's own construction left implicit, and proved the sqrt-map isometry, the closed-form distance, the constant curvature, and the local relationship to the JSD; Chapter 18 used this geometry to read the companion volume's own twelve-projection pool as an explicit configuration of points on a curved statistical manifold, exhibiting an exact coincidence (Proposition 18.2) and explicit, verified geodesics between them; this chapter closed the loop back to Part II by showing, through direct computation rather than appeal to Theorem 17.10's local theory alone, that Theorem 14.2's specific regime-wise minimizers are unchanged when the comparison functional is switched from JSD to Fisher–Rao distance (Proposition 19.1), while the two divergences' full rankings of the pool nonetheless disagree at two isolated, near-tied, non-extremal ranks (Section 19.4) — a nuanced, honestly reported answer rather than either a clean confirmation or a clean refutation of metric-independence. Chapter 21 now draws together all three parts of the monograph.

Chapter 20

Regional Pairwise Structure and Hierarchical Clustering of the Projection Pool

Chapter 18's closing section (Section 18.6) left open, explicitly, "a systematic account of the pairwise-distance-matrix structure at the three regional regimes" (only the Global-regime matrix was reported there, Table 18.1) and "any formal clustering analysis ...of the pairwise distance matrix." This chapter closes both items directly: it computes the full pairwise Fisher–Rao matrix at all four regimes of this monograph's own pipeline, and applies average-linkage hierarchical clustering to each.

20.1 Method

The pairwise Fisher–Rao distance $d_{\mathrm{FR}}(P_i, P_j)$ between every pair of the twelve pool projections is computed exactly as in Chapter 18, from the same ε-regularized, 60-bin pairwise-distance histograms already constructed by this monograph's own pipeline (Chapter 19), at each of the four regimes in turn — not only Global. Each regime's 12×12 distance matrix is condensed and passed to average-linkage (UPGMA) agglomerative clustering; the resulting dendrogram's cophenetic distances are compared against the original pairwise distances via the cophenetic correlation coefficient, a standard diagnostic of how faithfully a dendrogram represents the underlying distance matrix it was built from.

Why average linkage, and what this choice does not claim

Average linkage is used here as a standard, unremarkable default (not tuned or selected to produce a particular outcome); no claim is made that it is the uniquely correct or best-performing linkage rule for this data, only that it is a legitimate, widely used one. Section 20.6 returns to this point.

20.2 Cophenetic fidelity by regime

Table 20.1: Cophenetic correlation coefficient (fidelity of the average-linkage dendrogram to the underlying 12×12 pairwise Fisher–Rao matrix), by regime.

Regime	Cophenetic correlation
Global	0.9901
Regional mid-latitude	0.9193
Regional equatorial	0.9254
Regional high-latitude	0.8455

The Global regime's dendrogram is a highly faithful (correlation 0.99) summary of its own pairwise matrix; the three regional regimes are noticeably less faithful (correlation 0.85–0.93), consistent with a less strongly hierarchical structure among these twelve projections' behaviour on smaller, more geographically constrained grids, where the earlier-noted collapse in absolute IDI magnitude (Section 15.3) leaves less room for a clean nested structure to emerge.

20.3 Merge structure by regime

Table 20.2: First three merges (smallest Fisher–Rao distances, in bits) of the average-linkage dendrogram at each regime.

Regime	1st merge (height)	2nd merge (height)	3rd merge (height)
Global	Mercator–LCC (0.0000)	Aitoff–Hammer (0.0526)	+Mollweide (0.0809)
Regional mid-lat.	LambertAzEA–Albers (0.0692)	+LCC (0.0830)	Mollweide–Hammer (0.1016)
Regional equatorial	Mercator–LCC (0.0000)	Aitoff–Hammer (0.0219)	+Sinusoidal (0.0373)
Regional high-lat.	LCC–Albers (0.0351)	+LambertAzEA (0.0400)	Mollweide–Hammer (0.1106)

Two structural patterns recur across all four regimes, both consistent with, and now sharpened beyond, observations already made in passing elsewhere in this monograph:

1. **The conformal/equal-area conic family clusters tightly.** Lambert Conformal Conic, Lambert Azimuthal Equal-Area, and Albers Equal-Area Conic — three projections sharing a fitted-parallel or fitted-centre construction keyed to each grid's own latitude range — merge with each other before merging with any other projection, at every one of the three regional regimes (mid-latitude: steps 1–2; high-latitude: steps 1–2; equatorial: this trio does not cluster as tightly, merging only via the separate Lambert-Azimuthal/Sinusoidal/Aitoff/Hammer branch, itself notable as a regime-specific exception). At the Global regime, Mercator and Lambert Conformal Conic's exact zero-distance identity (Proposition 18.2) dominates the first merge instead.

2. **The whole-sphere compromise projections (Aitoff, Hammer, Mollweide, Winkel Tripel) cluster together, separately from the conic family.** This recurs at the Global and Regional high-latitude regimes explicitly (Table 20.2's 2nd/3rd-merge column and its high-latitude analogue in the full dendrogram) and, in a modified form (Aitoff and Hammer merging with Sinusoidal rather than Mollweide first), at the Regional equatorial regime. This is a direct, computed confirmation — via an entirely different method (hierarchical clustering of pairwise Fisher–Rao distances) than the JSD-based regime-wise ranking already used in Chapter 15 — of the same qualitative grouping (compromise azimuthal/pseudocylindrical projections behaving similarly to each other, and differently from the fitted-parallel conic family) that this monograph's Part II impossibility argument already exploited without formally clustering it.

Key Result

The pairwise Fisher–Rao structure of the twelve-projection pool is not merely a table of numbers (Table 18.1 and its three regional analogues, now computed in full for the first time in Table 20.2): it supports a genuine, four-regime-consistent hierarchical grouping into a fitted-parameter conic family and a whole-sphere compromise family, recovered independently by an unsupervised clustering method rather than assumed from the projections' design intent.

20.4 Linkage sensitivity

Section 20.1's disclaimer flagged average linkage as an unremarkable default, not a claim of uniqueness. This section checks that claim directly, re-running the same four distance matrices through single linkage and complete linkage as well, and comparing the resulting three-cluster partitions (via the Adjusted Rand Index, ARI, which is 1 for identical partitions and approximately 0 for agreement no better than chance).

Table 20.3: Pairwise agreement (ARI) between three-cluster partitions from single, complete, and average linkage, by regime. ARI $= 1$ indicates the two linkage rules produce identical partitions.

Regime	ARI(single,avg.)	ARI(complete,avg.)	ARI(single,complete)
Global	1.0000	1.0000	1.0000
Regional mid-latitude	0.1473	1.0000	0.1473
Regional equatorial	0.6992	1.0000	0.6992
Regional high-latitude	1.0000	1.0000	1.0000

Complete linkage agrees exactly with average linkage (ARI $= 1$) at all four regimes: the grouping reported in Section 20.3 is not an artefact of the specific choice between these two rules. Single linkage agrees exactly with the other two at the Global and Regional high-latitude regimes, agrees only partially (ARI 0.70) at the Regional equatorial regime, and disagrees close to chance level (ARI 0.15) at the Regional mid-latitude regime. Inspecting the mid-latitude single-linkage partition directly explains why: single linkage's well-known chaining tendency (a new point joins a cluster as soon as it is close to *any one* existing member, rather than close to the cluster as a whole) merges nine of the twelve projections into one large chained cluster at this regime, rather than resolving the conic-family/compromise-family split that complete and average linkage both recover.

Key Result

The conic-family/compromise-family grouping of Section 20.3 is robust to the choice between average and complete linkage at all four regimes, and to single linkage as well at the Global and Regional high-latitude regimes; it is specifically single linkage's chaining behaviour, not the underlying distance structure, that produces a different partition at the Regional mid-latitude and (more mildly) Regional equatorial regimes.

20.5 Cluster stability under grid resampling

Section 20.6's first draft flagged an entirely unchecked question: are the three-cluster partitions of Section 20.3 an artefact of the specific finite grid each regime happens to use, or would a differently sampled grid over the same geographic region recover essentially the same grouping? This section checks that directly, by subsampling.

20.5.1 Method

For each regime, 50 independent subsamples of 80% of that regime's own grid points (drawn without replacement) are formed; for each subsample, the pool histograms, pairwise Fisher–Rao matrix, and average-linkage 3-cluster partition are recomputed from scratch using exactly

the same routines as the full-grid pipeline, with one fixed methodological choice held constant across all replicates: the histogram bin edges are always the full grid's own edges (recomputing bin edges from each subsample's own smaller maximum pairwise distance would conflate genuine clustering instability with an artefact of shifting bin boundaries). Each replicate's 3-cluster partition is compared against the full-grid partition via the Adjusted Rand Index (ARI).

20.5.2 Results

Table 20.4: Cluster stability under 80% grid resampling, 50 replicates per regime. ARI $= 1$ indicates a replicate's partition exactly matches the full-grid partition; the final column is the fraction of replicates clearing a conventional ARI ≥ 0.8 "strong agreement" threshold.

Regime	Mean ARI	Std. ARI	Min. ARI	Frac. ARI ≥ 0.8
Global	0.910	0.202	0.239	0.80
Regional mid-latitude	0.977	0.127	0.147	0.96
Regional equatorial	0.994	0.042	0.699	0.98
Regional high-latitude	0.829	0.250	0.324	0.68

All four regimes show high mean ARI (0.83–0.99): the conic-family/compromise-family grouping of Section 20.3 is, on average, recovered under 80% resampling, not an artefact of the exact finite grid used. The ordering across regimes is nonetheless the opposite of what Section 20.2's cophenetic-correlation ordering might suggest: Regional equatorial, whose cophenetic correlation (Table 20.1) was middling (0.925), has by far the most stable partition under resampling (mean ARI 0.994, minimum 0.699 across all 50 replicates), while Global, whose cophenetic correlation was the highest of the four (0.990), shows the second-lowest resampling stability (mean ARI 0.910, and a full 20% of replicates falling below the 0.8 threshold). Cophenetic correlation measures how faithfully a dendrogram represents its own matrix; it does not measure how much that matrix, or the partition drawn from it, changes under resampling, and Table 20.4 shows these are genuinely different properties of the data, not two views of the same one.

Regional high-latitude is the least stable of the four regimes (mean ARI 0.829, only 68% of replicates above the 0.8 threshold), consistent with Section 20.4's independent finding that this regime's clustering signal is comparatively fragile — though there, fragility showed up as sensitivity to the *linkage rule* at fixed data, whereas here it shows up as sensitivity to the *data* at fixed linkage rule; the two are not the same failure mode, but both point to the same regime as the one where this monograph's clustering conclusions should be read with the most caution.

Key Result

The conic-family/compromise-family grouping recovered in Section 20.3 is stable, on average, to 80% grid resampling at all four regimes (mean ARI 0.83–0.99), directly closing the resampling-stability item left open at the end of an earlier draft of this chapter. Stability is not uniform, however: Regional high-latitude and, more intermittently, Global both produce a non-trivial fraction of low-ARI replicates, so the grouping should be read as a real but not perfectly rigid feature of the data at these two regimes specifically.

20.6 What remains open

Two questions are left explicitly open:

1. **Why cophenetic fidelity and resampling stability are not the same ranking.** Section 20.5 found the Global regime has the highest cophenetic correlation but only the third-highest resampling stability of the four regimes, and Regional equatorial the reverse; no structural explanation for this crossover is offered here beyond reporting the two rankings side by side.
2. **The Regional equatorial and high-latitude regimes' remaining exceptions.** Section 20.3 noted the conic family clusters less tightly at the equatorial regime than elsewhere, and Sections 20.4 and 20.5 both flag Regional high-latitude, by two independent measures (linkage-rule sensitivity and resampling stability), as the regime where the clustering signal is weakest; no explanation for why these two regimes specifically, rather than Global or mid-latitude, are the ones where the clustering signal is comparatively fragile is offered here beyond the observations themselves.

Chapter 21

Synthesis: Three Lenses on One Phenomenon

Section 1.1 opened this monograph with three theoretical questions, deliberately left independent of one another: how large must distortion provably be (Part I); can the Informational Distortion Index itself be beaten uniformly (Part II); and does the space of projections carry a natural, curved geometry beyond the JSD's bounded scalar summary (Part III). This closing chapter does not attempt to unify the three parts into a single theorem — Section 1.1 was explicit that no such unification was intended — but it does draw together, in one place, what each part actually established, where their boundaries of proof fall, and the handful of places where a method or a number from one part turned out, honestly rather than by design, to echo another.

21.1 What was proved, part by part

21.1.1 A stopping rule for Part I's azimuthal-profile sequence

Chapters 5 to 11 form a visible chain, each chapter opening by naming the specific open point of its predecessor that it closes. This is worth commenting on directly, since the same author's toponomastic series was closed, in its own final volume, on the explicit ground that continuing to extend it risked an unbounded sequence of ever-narrower follow-on pools with diminishing returns. Part I's chain could, in principle, have continued in exactly that way: each of Chapters 8 to 11 closed a point genuinely left open by its predecessor, and nothing in that pattern is self-limiting.

What stops it here is not fatigue with the pattern but a change in its kind, visible in Chapter 12's own closing section: the three sub-problems that chapter decomposes the remaining question into (symmetrization via a rearrangement inequality with no known analogue for a sup / inf-of-ratios functional; a global variational argument in the absence of any established convexity of $D_\rho(x)$ in ρ; existence of a minimizer over a non-compact-looking function space) each require mathematical machinery — rearrangement inequalities, minimax duality or Hamilton–Jacobi–Bellman-type arguments, direct-method existence theory — genuinely different in kind from the elementary, self-contained techniques (mean value theorem, convexity/concavity of a single real function, first-variation and envelope arguments) that carried every closure from Chapter 5 through Chapter 11. The chain of chapters was, in a precise sense, extensible only as long as each new open point could be closed with the same toolkit as the last; Chapter 12 is the chapter that found the toolkit's edge and said so, rather than forcing one further elementary-looking chapter onto a problem that no longer is one. Part I's azimuthal-profile investigation is accordingly treated, from this monograph's closure onward, as complete: the Global Embedding Optimality Problem of Definition 12.1 remains a genuine, precisely stated open problem, but one for a different investigation, using different tools, rather than the next chapter in this one.

Table 21.1: A consolidated inventory of this monograph's main results, and the explicit boundary between what is proved and what is left open in each case.

Part	Proved in full	Left explicitly open
I	Chordal benchmark for orthographic projection, distortion exactly $\sec(r/R)$, proved first on radial pairs alone (Theorem 3.3) and then, via the general curvature theorem, over *every* pair in the cap (Theorem 11.3); universal lower bound $D_{\min}(r)$ over *all* embeddings via Ptolemy's inequality (Theorem 4.3); exact closed-form global distortion $D_{\text{eq}}(r) = (r/R)/\sin(r/R)$ of the azimuthal equidistant projection (Proposition 5.2 and Theorem 5.6); a single unconditional curvature theorem covering any azimuthal profile of fixed curvature sign (Theorem 10.1), giving closed-form global distortion $D_{\text{ea}}(r) = D_{\text{st}}(r) = \sec^2(r/2R)$ for the equal-area and stereographic projections and $D_{\text{gn}}(r) = \sec^2(r/R)$ for the gnomonic projection, and confirming the azimuthal equidistant projection as strictly the best of all five classical azimuthal projections compared; local optimality of the equidistant profile along *every* admissible smooth or piecewise-smooth perturbation direction, with strict first-order decrease in every nontrivial direction (Theorems 8.8 and 9.2)	The Global Embedding Optimality Problem (GEOP, Definition 12.1): whether $D_{\text{eq}}(r) = D^{\star}(r)$ over *all* embeddings, not only azimuthal or perturbative ones — explicitly decomposed into three independent sub-problems (symmetrization, global variational analysis, existence of a minimizer), none attempted here (Chapter 12); the remaining factor-of-2 leading-order gap between $D_{\text{eq}}(r)$ and $D_{\min}(r)$; why the equal-area and stereographic projections' global distortion coincide exactly, left unexplained at the mechanism level (Section 6.5)
II	Impossibility of a uniform IDI minimizer, for the companion volume's own twelve-projection pool and four regimes (Theorem 14.2), by direct, executed enumeration; persistence of the same impossibility for a seventeen-projection pool and a fifth (Arctic) regime, with a *third* distinct regime-wise winner emerging (Chapter 16)	Impossibility for still larger or differently constructed pools/regimes; a fully universal (pool- and regime-free) impossibility theorem (Sections 14.4 and 16.4)
III	Fisher information metric of the categorical model and its sqrt-map isometry to a round sphere (Theorem 17.6); exact Fisher–Rao distance and constant curvature $1/4$ (Corollaries 17.8 and 17.9); local quadratic relation to the JSD (Theorem 17.10); metric-independence of Theorem 14.2's regime-wise minimizers (Proposition 19.1); full pairwise Fisher–Rao matrices and hierarchical clustering at all four regimes, recovering a consistent conic-family/whole-sphere-compromise grouping, checked and found robust across average, complete, and (at two of four regimes) single linkage, and stable on average (0.83–0.99 mean ARI) under 80% grid resampling at all four regimes (Chapter 20 and Sections 20.4 and 20.5)	Global JSD–Hellinger sandwich (numerically verified only, Proposition 17.12); geometry at the simplex boundary; any general account of when two projections' histograms coincide exactly (only the one instance of Proposition 18.2 is explained); why cophenetic fidelity and resampling stability rank the four regimes differently, and why Regional equatorial and high-latitude specifically are the regimes where the clustering signal is comparatively fragile (Section 20.6)

21.2 Three unplanned echoes across the parts

None of the following three observations was built into the monograph's design in advance — each surfaced only once the relevant chapter's computation was actually carried out — and each is reported here exactly as it was found, not as evidence of some deeper unity the three parts were secretly pursuing all along.

The sphere appears twice, playing structurally analogous roles. Part I's whole programme concerns the geographic sphere S^2_R, on which the impossibility of isometric embedding into the plane (Theorema Egregium) forces the distortion bounds of Chapters 3 and 4. Part III's Fisher–Rao geometry, developed for an entirely different purpose — comparing histograms, not geographic points — turns out (Theorem 17.6) to also be, exactly, a sphere: the statistical manifold of 60-bin categorical distributions is isometric to (an open subset of) a round sphere of radius 2, with its own constant positive curvature $1/4$ (Corollary 17.9), entirely unrelated numerically to the geographic sphere's curvature $1/R^2 \approx 1/6371^2\,\mathrm{km}^{-2}$ but structurally the same kind of object. Two constant-curvature spheres, of utterly different provenance and scale, underlie the two halves of this monograph's mathematics.

A local linearization recurs as the bridge between an exact pointwise theory and a global one. Remark 2.2 showed that Tissot's classical indicatrix is exactly the small-neighbourhood limit of Definition 2.1's two-point distortion; Theorem 17.10 showed, by an entirely independent Taylor-expansion argument in a completely different space (the statistical simplex, not the geographic sphere), that the JSD is exactly the small-neighbourhood limit, to leading order, of the Fisher–Rao distance's squared value. In both cases, a locally linear (Part I: differential; Part III: quadratic-form) approximation to a genuinely nonlinear, global quantity turns out to be exact only infinitesimally, with the gap between local and global measured, and left open, explicitly (Section 3.5's non-radial extension; Proposition 17.12's unproved global sandwich).

The same adversarial pair anchors both Part II's impossibility theorem and Part III's cross-validation of it. Aitoff and Lambert Azimuthal Equal-Area are singled out, independently, by Theorem 14.2's proof (as the two regime-wise JSD minimizers whose trade-off across regimes drives the impossibility) and by Proposition 19.1's Fisher–Rao recomputation (as the same two regime-wise minimizers, unchanged under a structurally different divergence). This is not a coincidence built into the monograph's design — Part III's pipeline was written and executed without adjusting which projections or regimes to test, precisely so that an agreement or disagreement with Part II's already-published result would be a genuine finding rather than a foregone conclusion (Section 19.3) — and the agreement found is itself the most direct substantive link between the two parts.

Extending the data, in both Part II and Part III, sharpened rather than softened the original conclusions. Chapter 16's extension of Part II from twelve to seventeen projections and four to five regimes did not produce a new uniform minimizer approaching consensus; it produced a *third* distinct regime-wise winner, making the impossibility more visible, not less. Chapter 20's extension of Part III's pairwise Fisher–Rao analysis from one regime to all four did not contradict the qualitative conic-family/compromise-family grouping that Part II's own argument (Section 14.3's plausibility account) had already informally invoked to explain *why* the impossibility takes the specific shape it does; it recovered that same grouping independently, by an unsupervised method applied to a different divergence, at three further regimes. Neither extension was guaranteed to come out this way — a larger pool could have produced a consensus winner, and clustering could have found a grouping unrelated to the projections' known construction — and both are reported here exactly as executed, not selected after the fact for agreeing with the smaller-scale results.

21.3 What this monograph does not claim

Consistent with the discipline stated at the outset (Section 1.2) and maintained, chapter by chapter, throughout (every chapter of this monograph has closed with an explicit "what remains open" section), it is worth restating plainly, in one place, what the three parts together do *not* establish. This monograph does not prove a fully general, pool- and embedding-independent impossibility theorem for cartographic distortion of any kind: Part I's universal lower bound (Theorem 4.3) is universal over embeddings but specific to one spherical-cap configuration and, even after Chapter 5's narrowing, leaves a factor-of-2 sharpness gap open; Part II's impossibility theorem (Theorem 14.2) is specific to one finite projection pool and regime collection; Part III's geometry is specific to one choice of statistical model (the 60-bin categorical histogram) among the several a different author might have chosen. Each part's contribution is real, checked, and in the cases marked "proved" in Table 21.1 fully rigorous — but the three parts are three genuinely partial illuminations of the cartographic distortion phenomenon, not three facets of one completed theory, and this monograph has tried throughout to make the edges of each illumination visible rather than blurring them into a false appearance of completeness.

21.4 Closing remark

The companion volume's empirical programme found, repeatedly and across a growing pool of projections and regions, that no single map projection is best for every purpose. This monograph set out to ask, of that empirical pattern, three different kinds of theoretical question — how much distortion is forced, whether the pattern can be proved rather than merely observed, and what geometry the underlying comparison itself carries — and found, in each case, a real but bounded answer: a benchmark distortion bound, not yet a sharp one; a proof of impossibility, for one pool and regime collection, not a fully universal law; a curved statistical geometry, made completely explicit for one specific statistical model, whose relationship to the JSD is proved locally and only well supported — not proved — globally. If there is a single thread connecting all three, it is this: in each part, the honest, checkable content of the result is narrower than the informal empirical observation that motivated it, and stating that narrower content precisely, rather than the broader claim it merely suggests, has been the organizing discipline of this monograph throughout.

References

This monograph draws on classical results in metric embedding theory, comparison geometry, information geometry, and map projection design, alongside three pieces of notation and one computational pipeline reused without modification from the companion volume, *Original Mathematical Perspectives on Cartography* (cited here as a whole; individual chapter numbers are given in the text at each point of reuse).

Bibliography

[Amari & Nagaoka, 2000] Amari, S., & Nagaoka, H. (2000). *Methods of Information Geometry*. Translations of Mathematical Monographs, Vol. 191. Providence, RI: American Mathematical Society.

[Assouad, 1983] Assouad, P. (1983). Plongements lipschitziens dans $\mathbb{R}^n$. *Bulletin de la Société Mathématique de France*, 111, 429–448.

[Bourgain, 1985] Bourgain, J. (1985). On Lipschitz embedding of finite metric spaces in Hilbert space. *Israel Journal of Mathematics*, 52(1–2), 46–52.

[Foertsch, Lytchak & Schroeder, 2007] Foertsch, T., Lytchak, A., & Schroeder, V. (2007). Nonpositive curvature and the Ptolemy inequality. *International Mathematics Research Notices*, 2007, rnm100.

[Danskin, 1967] Danskin, J. M. (1967). *The Theory of Max-Min and its Application to Weapons Allocation Problems*. Ökonometrie und Unternehmensforschung, Vol. 5. Berlin: Springer-Verlag.

[Gauss, 1828] Gauss, C. F. (1828). Disquisitiones generales circa superficies curvas. *Commentationes Societatis Regiae Scientiarum Gottingensis Recentiores*, Vol. 6, 99–146.

[Kass, 1989] Kass, R. E. (1989). The geometry of asymptotic inference. *Statistical Science*, 4(3), 188–219.

[Lin, 1991] Lin, J. (1991). Divergence measures based on the Shannon entropy. *IEEE Transactions on Information Theory*, 37(1), 145–151.

[Linial, London & Rabinovich, 1995] Linial, N., London, E., & Rabinovich, Y. (1995). The geometry of graphs and some of its algorithmic applications. *Combinatorica*, 15(2), 215–245.

[Rao, 1945] Rao, C. R. (1945). Information and the accuracy attainable in the estimation of statistical parameters. *Bulletin of the Calcutta Mathematical Society*, 37, 81–91.

[Snyder, 1993] Snyder, J. P. (1993). *Flattening the Earth: Two Thousand Years of Map Projections*. Chicago: University of Chicago Press.

[Snyder & Voxland, 1989] Snyder, J. P., & Voxland, P. M. (1989). *An Album of Map Projections*. U.S. Geological Survey Professional Paper 1453. Washington, DC: U.S. Government Printing Office.

[Wolpert & Macready, 1997] Wolpert, D. H., & Macready, W. G. (1997). No free lunch theorems for optimization. *IEEE Transactions on Evolutionary Computation*, 1(1), 67–82.